"Essays on the Efficient Use of Public Resources"

Dissertation
for the Faculty of Economics, Business Administration and Information Technology of the University of Zurich

to achieve the title of
Doctor of Philosophy
in Economics

presented by

Philippe K. Widmer
from Reiden (LU)

approved in April 2012 at the request of

Prof. Dr. Peter Zweifel
Prof. Dr. Dieter Pfaff

The Faculty of Economics, Business Administration and Information Technology of the University of Zurich hereby authorises the printing of this Doctoral Thesis, without thereby giving any opinion on the views contained therein.

Zurich, 04. April 2012

Chairman of the Doctoral Committee: Prof. Dr. Dieter Pfaff

Preface

This doctoral thesis was written during my time as a research assistant at the Department of Economics (formerly the Socioeconomic Institute) of the University of Zurich, Switzerland, from 2007 to 2011. The time I spent at the University of Zurich was one of great inspiration and was full of new insights. I have enjoyed the help and support of several excellent people whom I would like to thank.

First and foremost, I am grateful to Prof. Dr. Peter Zweifel, my thesis supervisor, for finding interest in my research topics and for giving considerable freedom to write this doctoral thesis. His guidance, support, and intellectual inputs have been greatly inspiring and most valuable. I would also like to thank him for giving me responsibility for several interesting empirical projects. Furthermore, I want to thank Prof. Dr. Dieter Pfaff, who kindly consented to be my second supervisor.

In addition, I would like to thank my former and current work colleagues at the Department of Economics: George Elias, Prof. Dr. Mehdi Farsi, Dr. Patrick Eugster, Dr. Boris Krey, Dr. Ilja Neustadt, Maurus Rischatsch, Adrian Rohner, Micha Ruflin, PD Dr. Felix Schläpfer, Dr. Michèle Sennhauser, Dr. Harry Telser, and Maria Trottmann. Special thanks go to Philipp Morf, and Dr. Johannes Schoder. I really enjoyed the countless discussions about economics, research, and "..." that not only shaped the ideas of this thesis. Moreover, I am also particularly grateful to the librarians of the Department of Economics who were always willing and happy to help.

I am also indebted to my colleagues at the Polynomics AG for all the experiences in daily business and for the many challenging questions to the mechanics of efficiency measurement. This experience was of high value during the writing of this thesis.

This is also to thank my parents Karl and Margrith Widmer for their unconditional support throughout my whole life. Their encouragement and care helped me enormously in reaching my goals. Furthermore, to my brothers and sisters Matthias and Claudia for all their patience and motivations throughout this journey.

Most of all, I want to thank my partner in love, Nathalie Wismer for her invaluable support and encouragement during my PhD studies. During times of great challenges she always managed to make me smile and to carry on. Her belief in my capabilities enabled me to complete this thesis.

Philippe Widmer
Zurich, September 2011

Contents

List of Figures

List of Tables

CHAPTER I

Introduction: Theme, Objectives, and Structure

1

Introduction: Theme, Objectives, and Structure

The first theorem of welfare states that under perfect market conditions any economy should achieve allocative and productive efficiency, also known as Pareto efficiency (Pareto, 1897). Under Pareto efficiency, no one can be made better off without making someone else worse off and the government only needs to engage in some initial redistributions through lump-sum transfers. However, the Greenwald-Stiglitz theorem states that the conditions for Pareto efficiency seldom exist. Market failures occur almost everywhere, suggesting that the government could improve welfare by resource allocation and efficiency gains (Stiglitz, 1991).

So far, government interventions have typically been limited to resource allocation, ignoring the potential for efficiency gains. As the debt crises of the United States and Europe demonstrate, this does not necessarily increase welfare. In contrast, welfare could even decrease if resource allocation creates incentive to waste public resources. Yet, since demographic changes and globalization efforts continue to limit the public resources in many developed countries, production efficiency must be addressed.

This dissertation focuses on efficiency measurements and discusses four topics where government interventions could cause incentive for market inefficiencies. The countries analyzed are Switzerland and the United States. In all chapters, advanced non-parametric and parametric efficiency measurement approaches are applied to deal with several important issues, such as the relevance of unobserved heterogeneity and uncertainty in production inputs and outputs.

Chapter 2 focuses on the efficiency of public good provision in Switzerland. From a theoretical point of view, Tiebout competition should induce Swiss member states (cantons) to use tax revenue efficiently (Tiebout, 1956). However, this requires that there are no externalities and that differences in performance are due entirely to the efficiency of administration. Since disparities must be mitigated in order to create a level playing field, Switzerland has applied a fiscal equalization scheme that enables poor cantons to produce public goods at average tax rates. However, despite the aspiration for equity gains, fiscal equalization can increase disparities if cantons on the receiving end lack incentives for efficiency, commonly known as the "flypaper effect" (Inman, 2008). The efficiency of contributing cantons may be undermined as well, giving rise to an equity-efficiency trade-off (Stiglitz, 1988).

This chapter includes an investigation of whether fiscal equalization among Swiss cantons reduces the incentive for efficiency. The efficiency of all 26 Swiss cantons is measured between 2000 and 2004. Aggregate output performance indicators, including six major public services, are constructed to calculate cantonal efficiency scores based on robust data envelopment analysis. Efficiency scores are then related to the fiscal equalization scheme operated by the Swiss federal state, controlling for the socioeconomic factors that also influence cantonal performance.

Chapter 3 and 4 provide an empirical analysis of the efficiency of Swiss hospitals. As in many other developed countries, increasing health care expenditures have highlighted the importance of health care reforms that pursue efficiency gains. One such reform includes transitioning to a prospective payment system. The assumption is that a change to predetermined and fixed payments places hospitals at operating risk and increases their cost efficiency (e.g., Biorn et al., 2003, Chalkley and Malcomson, 1998, and Newhouse, 1996). However, the challenge for policy makers is how to reward efficiency, which involves identifying the differences caused by inefficiency instead of by heterogeneity due to exogenous influences. If regulators enforce the cost reductions indicated by simple performance measures, such as operating cost per casemix-adjusted patient case, highly efficient hospitals could end up in financial distress. Chapter 3 includes an empirical analysis of the importance of heterogeneity in the measurement of Swiss hospital efficiency. Particularly in federalist countries such as Switzerland, where hospitals operate in different regulatory environments and provide health care services using different technologies, heterogeneity may significantly influence the cost variability among hospitals. In Chapter 3, stochastic frontier analysis is applied to a standard, random intercept, and random parameter frontier model in order to account for heterogeneity in the production technology. Estimates are derived from a variable cost function for approximately 100 Swiss hospitals for the years 2004 to 2007.

Since it is difficult to account for heterogeneity when estimating hospital cost efficiency, the question remains whether prospective payment systems can contain hospital costs. The purpose of Chapter 4 is to analyze the influence of prospective payment systems on Swiss hospital efficiency. The analysis provides a comparison of a retrospective per diem payment system with a prospective global budget and a payment per patient case system. Again, this chapter includes a stochastic frontier analysis using a standard and a random parameter frontier model to account for the existence of heterogeneity. Estimates are derived using a variable cost function for approximately 90 subsidized Swiss hospitals for the years 2004 to 2009.

An analysis of the relevance of efficiency measurement in the provision of electricity is presented in Chapter 5. Stochastic frontier analysis and data envelopment analysis are the two most prevalent approaches for the measurement of efficiency (see Chapter 2 to 4). However, they are only suitable when productive units are homogenous with regard to technology (relaxed in Chapter 3 and 4) and have stable input prices, and hence little uncertainty. In the provision of electricity, both of these assumptions are not satisfied. In particular, each power plant utilizes a different type of technology depending on its primary energy source (e.g., coal, nuclear, wind) and is exposed to different exogenous shocks (e.g., the Gulf war in the case of oil). Exogenous shocks cause unexpected changes in input prices that affect the level and development of the efficient frontier. Therefore, it is not sufficient to focus on the lowest cost provision of electricity. A risk-adjusted efficiency measurement that involves the optimal mix of technology provides a more reliable measurement. In Chapter 5, portfolio theory is applied to estimate the efficiency of electricity provision in the United States and Switzerland. Seemingly unrelated regression estimation (SURE) is adopted for estimating the covariance matrix used in determining the efficient frontier.

Chapter 6 concludes by stating the major policy implications and disclosing possible future extensions.

Note that Prof. Dr. Peter Zweifel co-authored Chapters 2, 3 and 5, Prof. Dr. Mehdi Farsi co-authored Chapter 3, and Dr. Boris Krey co-authored Chapter 5. Chapter 2 is forthcoming in *Public Finance Review*. Chapter 3 has been submitted to the *Journal of Productivity Analysis*. Chapter 4 has been submitted to *Health Economics*, and Chapter 5 has been submitted to *Energy Policy*.

This introduction concludes with a note concerning the structure of this dissertation. Each chapter of this dissertation can be considered as self-contained, having its own appendix. Institutional features are explained separately in each chapter in the interest of readability. References across chapters are made explicit.

CHAPTER II

Fiscal Equalization, Tiebout Competition, and Incentives for Efficiency in a Federalist Country

Philippe K. Widmer and Peter Zweifel

Abstract: The purpose of this paper is twofold. First, it measures the efficiency in the provision of public goods by local jurisdictions applying Data Envelopment Analysis (DEA). Second, it relates efficiency scores to a fiscal equalization scheme designed to mitigate the negative consequences of Tiebout competition. The data come from the 26 cantons of Switzerland (2000-2004), a country characterized by marked federalism. Results show the equalization scheme to indeed have a negative influence on performance, resulting in an efficiency-equity trade-off (Stiglitz, 1988). However, substitution of earmarked payments by lump-sum payments as part of the 2008 reform is likely to enhance cantonal performance.

Keywords: DEA, efficiency measurement, federalism, fiscal equalization, public finance, Switzerland, Tiebout competition

JEL classification: C14; C67; H11; H72; H83

2

Fiscal Equalization, Tiebout Competition, and Incentives for Efficiency in a Federalist Country

2.1 Introduction

During the past decade, growing tax burdens have combined with ecological and equity concerns to increase citizens' interest in the efficient provision of public goods. Economists have been responding to this interest by trying to provide information about government performance that may contribute to an efficient use of tax revenues. Examples of efficiency measurement of public services include Drake and Simper (2003), who examined police departments in England and Welsh, Worthington and Dollery (2001), who estimated the efficiency of waste management in South Wales and Worthington (2001), who focused on U.S. and English public education. Grossman et al. (1999) conclude that competition between U.S. cities serves to increase their efficiency, in line with the Tiebout hypothesis. As to continental Europe, Afonso and Fernandes (2006); Afonso and Scaglioni (2005); De Borger and Kerstens (1996) as well as Vanden Eeckaut et al. (1993) examined the efficiency of Lisbon, Italian, and Belgian local governments, respectively. Specifically, De Borger and Kerstens (1996) find that the tax rate and income per capita have an insignificant effect on the performance of Belgian local governments, while federal grants have a negative influence. At the country level, Afonso et al. (2006), comparing new EU member and emerging market states, conclude that trade openness and transparency in government have a positive but insignificant effect on efficiency, while public trust in politicians fosters inefficiency.

These studies have not taken into account one feature of federalist countries that may affect efficiency at the local level, viz. fiscal equalization schemes. Fiscal equalization is designed to reduce horizontal and vertical fiscal imbalances that often exist between lower-level jurisdictions to provide public goods. This reduction is achieved by payments from jurisdictions with above-average fiscal capacity to jurisdictions with below-average fiscal capacity. In this way below-average jurisdictions are to be enabled to produce public goods at average tax rates (Thöny, 2005). Equalization schemes exist in most countries, among them the United States, the European Union, Germany, Austria, and Switzerland – sometimes even at the community level. However, little attention has been given to the influence of such programs on the performance of both contributing and receiving member states. Indeed, disparities in the provision of public goods could even increase because jurisdictions on the receiving end may lack incentives for efficiency,commonly known as the "flypaper effect" (Inman, 2008). The efficiency of contributing states may be undermined, too, giving rise to the well-documented equity-efficiency trade-off (Stiglitz, 1988).

The contribution of this paper therefore is twofold. First, it measures the efficiency of all 26 Swiss cantons between 2000 and 2004. Aggregate output performance indicators including six major public services are constructed to calculate cantonal efficiency scores based on robust Data Envelopment Analysis (DEA). Second, calculated efficiency scores are related to the fiscal equalization scheme operated by the Swiss federal state both in its present and its new (allegedly improved) form, controlling for socioeconomic factors that also have an influence on cantonal performance.

To the best knowledge of the authors, this is the first contribution undertaking a macroeconomic efficiency measurement of public good provision in a federalist country that takes the incentive effects of a fiscal equalization scheme into account.

This paper is organized as follows. The second section provides some background information about Swiss federalism. The third section contains a review of efficiency measurement methods to argue that DEA is the method of choice in the present context. The data used are described in the fourth section. The fifth section is devoted to the presentation of results of the DEA and of a Tobit model estimating the effect of the fiscal equalization scheme on DEA efficiency scores. The final section concludes with an outlook and suggestions for future research.

2.2 Swiss Federalism

2.2.1 Cantons as the Producers of Public Goods

Switzerland, a federal state with its constitution dating from 1848, distinguishes between three levels of government, viz. federal, 26 cantons[1], and approximately 2,600 communities. Public services are financed and provided at all three levels, but with differing authorities. While the communities act under cantonal oversight, the cantons still constitute the backbone of the state. By article 3 of the Swiss constitution, they are responsible for all public services that are delegated neither to the federal state nor to their affiliated local authorities. Cantons are sovereign governmental entities with their own constitution and separation of power (legislative, executive, and judiciary), resulting in an extremely decentralized provision of public services.

Table 2.1 shows public expenditure on the 12 major service categories according to the three levels of authority. To the extent that Olson's (1969) equivalence principle applies, expenditure by an authority also means provision. According to that principle, more than 60 percent of public good provision are estimated to be controlled by the 26 cantons and their affiliated communities. However, this share varies between categories; it is particularly low for military defense and foreign relations, which are delegated to the federal state. It is highest in education and health, which also constitute two of the most important overall expenditure items.

The Tiebout (1956) hypothesis predicts a positive relationship between fiscal federalism and performance of government. Similar to a free market economy, where consumers buy from the producer offering the best performance-price ratio, citizens choose the juris-

[1] The 26 Swiss cantons are Appenzell Inner-Rhodes (AI), Appenzell Outer-Rhodes (AR), Argovia (AG), Basel-City (BS), Basel-Country (BL), Bern (BE), Fribourg (FR), Geneva (GE), Glarus (GL), Grisons (GR), Jura (JU), Lucerne (LU), Neuchatel (NE), Nidwalden (NW), Obwalden (OW), Schaffhausen (SH), Schwyz (SZ), Solothurn (SO), St.Gall (SG), Thurgovia (TG), Ticino (TI), Uri (UR), Valais (VS), Vaud (VD), Zug (ZG), and Zurich (ZH).

Table 2.1: Functional Structure of Public Good Provision, Year 2004

Expenditure In CHF Million		Federal state	Cantons	Communities	Total
(1)	Administration	1,918	3,299	3,637	8,855
(2)	Public safety	728	5,287	1,955	7,970
(3)	[Military defense]	4,637	157	185	4,979
(4)	[Foreign relations]	2,427	-	-	2,427
(5)	Education	5,231	14,399	8,055	27'684
(6)	Culture & Sport	447	1,380	2,422	4,249
(7)	Health	200	12,203	6,922	19,326
(8)	[Social welfare]	13,805	8,026	5,911	27,742
(9)	Transportation	8,547	2,873	2,991	14,411
(10)	Environment & Spatial planning	728	1,019	3,159	4,907
(11)	Public economy	4,546	1,287	512	6,344
(12)	[Finance & Tax]	9,411	-984	1,059	9,486
Total expenditure		52,624	48,947	36,808	138,379

Source: Swiss Federal Statistical Office, 1 CHF = 0.8 USD (2004 exchange rates).

diction where they get the best ratio between public services provided and tax paid. In the case of Switzerland, cantonal autonomy in combination with direct democratic control through popular initiatives and referenda have resulted in considerable heterogeneity in the mode of provision. Since citizens can migrate and shift capital freely between cantons, they indeed expose them to Tiebout competition.

However, this hypothesis assumes that there are no externalities and that differences in performance are entirely due to the efficiency of administration. Externalities (spillover effects) exist if citizens from one canton cannot be prevented from using services provided by another canton without paying. They typically arise in health care, education, and culture, although cantons with specialized hospitals do charge higher fees to patients from elsewhere, those with a university levy higher tuitions, and those with an opera house often make other cantons contribute to their operating expense. As to the efficiency of administration, there are disparities that are due to topographic, demographic, and socioeconomic conditions, constituting a handicap that cannot be overcome by the affected canton. Both confounding influences will be controlled for (see Section 2.4.1 and 2.4.4, respectively) when assessing the influence of Tiebout competition on cantonal performance.

2.2.2 Existing Fiscal Equalization Scheme

To overcome these disadvantages of fiscal federalism, Switzerland initiated a fiscal equalization program in 1959 to equalize cantonal disparities in the provision of public goods. According to an amendment of the federal constitution (article 135), cantonal disparities are to be mitigated with reference to Tiebout competition, the objective is to create a level playing field. By 2004, fiscal equalization has grown to some 1,000 CHF mn. of payments from the confederation to the cantons and another 1,500 CHF mn. between them. In relative terms, it totals almost 3 percent of cantonal and communal expenditure. The program is geared to the 'financial potential' indicator, which has four components. Financial potential is defined to increase with

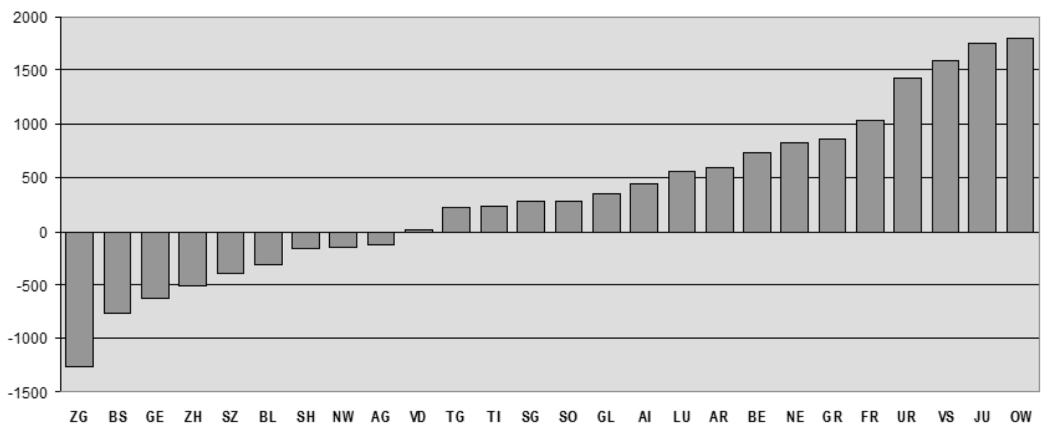
(1) *Income:* Cantonal income per capita;

(2) *Tax power:* Taxable income, weighted by the *tax burden* per capita;

(3) *Inverse of tax burden:* Cantonal plus communal taxation as a share of *Income*;

(4) *Favorable topographic situation:* Share of a canton's non-mountainous cropland in its total area, weighted by the number of inhabitants per unit of productive land.

Figure 2.1: Payments of the Swiss Fiscal Equalization Program (2004)$^{a)}$



Source: Federal Finance Administration (FFA)
a)For the acronyms, see footnote No. 1.

A higher total index value results in less financial assistance. Figure 2.1 shows total payments per capita as of 2004. The canton of Zug contributed the maximum of some

CHF 1,250 (1 CHF = 0.8 USD in 2004) per capita to the program, followed by Basel-City, Geneva, and Zurich. At the other extreme, the 33,000 inhabitants of the canton of Obwalden in central Switzerland received some CHF 1,800 on average. In comparison, the extreme values of the German equalization scheme are a maximum of some CHF 600 paid by the land of Hessen and a maximum of some CHF 1,200 CHF per capita received by Berlin. These figures illustrate the importance of the Swiss fiscal equalization scheme.

One also needs to distinguish between earmarked (almost 70 percent of total) and general payments. While general payments can be used by the canton in ways it believes to generate the highest benefit for its citizens, earmarked subsidies may result in gold plating of projects and hence inefficiency (see e.g. De Borger and Kerstens, 1996).

Wrong incentives of the fiscal equalization program could have a sizable influence on cantonal performance and national welfare. Indeed, the existing program has been suspected of inducing the disparities it is designed to alleviate. Especially components No. 2 and 3 of the index formula are seen to create incentives for subsidized cantons to keep their tax burden high, e.g. by using their tax revenue for projects that contribute little to economic growth but enhance politicians' popularity (Fischer et al., 2003). In addition, cantons that are obliged to pay into the scheme have incentives to waste their money as well. They rather spend it on idle projects than give it to other cantons. These concerns have resulted in a reform proposal that passed a popular referendum in 2006. Starting in 2008, the share of earmarked payments was to be reduced to a minimum. Equalization payments are to be governed by a new formula, which distinguishes between resource and financial disparities. Against this backdrop, this paper seeks to answer two questions:

(1) Does a fiscal equalization program as sizable as the Swiss contain incentives to provide public goods less efficiently, creating a trade-off between equity and efficiency?

(2) Does it matter whether transfer payments are earmarked or not?

2.3 Measuring Technical Efficiency with Data Envelopment Analysis

The characterization of Swiss cantons in the preceding section justifies viewing them as largely independent producers of a subset of public goods. For productivity measurement, they constitute decision making units (DMU) that transform inputs into outputs, with productivity reflecting the quality of their administration. Following Koopmans (1951), technical efficiency in the provision of public goods thus can be measured with reference to a technology set Γ,

$$\Gamma = \{(X, Y) |\; Y \leq f(X)\} \tag{2.1}$$

that describes the feasible set of input and output combinations (X, Y) of a production process. A DMU is called technically efficient if it lies on the boundary of Γ. On that boundary, it is not possible to produce more outputs Y for a given amount of inputs $\bar{X}$; or conversely, no smaller quantity of inputs X can produce a given output $\bar{Y}$.

There are various assumptions regarding the boundary of Γ. For simplicity we adopt those of Shephard (1970),

$$\begin{aligned} Iso\; X(y) &= \{x|\; x \in X(y), \theta x \notin X(y), \forall\; 0 < \theta < 1\} \\ Iso\; Y(x) &= \{y|\; y \in Y(x), \theta^{-1} y \notin Y(x), \forall\; 0 < \theta < 1\}. \end{aligned} \tag{2.2}$$

Here, the input and output isoquants $Iso\;(\cdot)$ define sections with strong and weak technical efficiency, depending on the slope of the frontier, with θ denoting a scalar by which all inputs can be reduced without leaving the feasibility set. Accordingly, θ^{-1} symbolizes the scaling-up factor for the outputs. However, the relevant technology set is almost never known in applied economic research, forcing the analyst to use observed rather than efficient input and output quantities. The pertinent methodology was developed by Farrell (1957); it has evolved into a distinction between parametric (econometric) and non-parametric (mathematical) methods (see Coelli et al., 2005 for respective overviews).

In public good provision analysis, Data Envelopment Analysis (DEA) is the most common alternative. Webster et. al (1998) argue that DEA dominates its main competitor, Stochastic Frontier Analysis (SFA) because of the following reasons:

- DEA is more flexible because no specific functional form of the transformation process needs to be specified;

- DEA does not have to rely on price data for inputs and outputs, which often is lacking in the public sector.

DEA is the preferred technique for the present investigation, in particular because of lacking information about factor prices. Public sector accounts are notorious for neglecting capital user cost, and Switzerland is no exception. The DEA version employed here is an input-orientated one. The objective is to determine an efficient frontier $\widehat{IsoX(y)}$ that is defined by the most productive DMUs. DEA amounts to solving a linear optimization problem for a particular DMU$_c$ or canton $c = 1, ..., 26$, with an $1 \times K$ output vector y_c and a $1 \times M$ input vector x_c,

$$\begin{aligned} \mathrm{Max}_{\upsilon,\nu} \quad & \upsilon' y_c \\ \text{s.t.} \quad & \nu' x_c = 1 \\ & \upsilon' Y - \nu' X \leq 0 \\ & \upsilon, \nu \geq 0. \end{aligned} \tag{2.3}$$

Here, the $26 \times K$ output matrix Y and $26 \times M$ input matrix X represent the data for all 26 cantons. Thus, let a canton optimize its outputs y_c and inputs x_c by maximizing the distance between them valued using weights υ and ν. Note that these weights relate to the universe of all cantons and can be interpreted as shadow prices. Moreover, inputs are normalized to sum up (after weighting) to 1. The inequality $\upsilon' Y - \nu' X \leq 0$ prevents outputs from increasing without bounds for a given bundle of inputs. $\widehat{IsoX(y)}$ is defined by those units for which $\upsilon' Y - \nu' X = 0$. Their efficiency score $\widehat{EFF_c}$ is 100 percent, while that of the other DMUs is given by their radial distance from the frontier.

However, the location of the efficient frontier strongly depends on the extreme DMUs (which lack comparators). One way to obtain robust DEA efficiency scores $\widehat{EFF}$ is to iteratively exclude one DMU lying on the efficiency frontier. The new frontier then assigns this DMU a so-called super-efficiency score in excess of 100 percent without truncating the scores of the remaining DMUs (Andersen and Petersen, 1993). The larger the super-efficiency of a DMU, the farther away it is from the remaining units in the technology set. Here, if this score is more than 1.5 times the distance between the 25th and the 75th percentile of all super-efficiency values, the pertinent DMU is excluded as an outlier (Thanssoulis, 1999).

In a second step, the obtained robust efficiency scores $\widehat{EFF_{ct}}$ (of cantons $c = 1, ..., 26$ in year $t = 2000, ..., 2004$) are related to a set of variables characterizing the Swiss fiscal equalization program and the disparities between cantons in order to address the two research questions stated at the end of Section 2.2,

$$\widehat{EFF_{ct}} = \gamma_0 + \gamma_1 Z_{1,ct} + \gamma_2 Z_{2,ct} + ... + \gamma_n Z_{n,ct} + \varphi_{ct}. \tag{2.4}$$

Commonly, Tobit estimation is applied to account for the fact that scores cannot exceed 1.00 (the lower limit of 0 is less relevant because it is never binding). Specification details are given in Section 2.5.2 below. Alternatively, one could also use the super-efficiency scores of the outlier detection process as dependent variable and estimate OLS. However, to be consistent with the theoretical background of Farrell (1957) and because the two variants lead to similar conclusions, we proceed with the more common robust efficiency scores as the dependent variable (see De Borger and Kerstens, 1996 and Drake and Simper, 2003 for applications of Tobit estimates).

2.4 Data

2.4.1 Service Categories Retained

The data come from the Federal Statistical Office, covering the years 2000 to 2004. As shown in Table 2.1, not all categories of services listed are predominantly subject to cantonal control. Moreover, the quality of data is insufficient for some categories. Therefore, only six out of twelve are retained for this investigation, viz. (1) administration, (2) public safety, (5) education, (7) health, (9) transportation and (11) public economy (they will be renumbered 1 to 6 below). Further, in order to exclude spillovers as far as possible, only primary and secondary education (without tertiary and vocational components), private road transportation (without regional public transportation) and farming and forestry are included in the the analysis. More refined adjustments for spillovers (known to exist especially in health care) were not possible. They are controlled for in the second step Tobit estimation.

2.4.2 Constructing an Aggregate Output Performance Index

Section 2.3 makes clear that the choice of output variables has an important influence on the results of a DEA. However, in public good provision, choice is no simple task because of two reasons. First, most of the outputs are not directly quantifiable. Second, public services are too many for individually entering them in a DEA. In this paper, we try to overcome these difficulties by running a cost driver analysis and by constructing an aggregate output performance index.

Selecting the Output Variables

Since outputs of the public sector are difficult to measure, activity-based indicators serve as a substitute, in line with previous studies (see e.g. Afonso et al., 2006). In this paper, two to six indicators for each of the six retained categories – 22 in total – are selected to proxy the output of public goods provided by a canton (see Table 2.2).

Our selection was based on two concerns: choosing the most relevant variables and making sure that they cover the years 2000-2004 for each canton. The relevance of the selected variables is checked with an analysis of cost drivers in the six service categories. Thus, the dependent variables are category-specific real expenditure C_{ct}^{j}, $(j = 1, ..., 6)$ of

Table 2.2: Output Indicators for the Six Governmental Activities Investigated

Public Service	Output	Description, remarks
(1) Administration		
Legislative, Executive	Y_1 Population	Population served and number of firms serve
General administration	Y_2 No. firms	as proxies for administration services provided.
(2) Public safety		
Judicature	Y_1 No. delinquencies	The assumption is that all citizens and
Police	Y_2 Population	dwelling units have the same preferences for
Fire department	Y_3 No. dwelling units	public safety and a similar probability of suffering from crime and fire risk.
(3) Education		
Kindergarden	Y_1 No. students	The numbers of kindergarten, primary,
Primary education	Y_2 No. students	secondary, and high school students serve as
Secondary education	Y_3 No. students	indicators of output values.
High school education	Y_4 No. students	
(4) Health		
Hospitals (specialized)	Y_1 No. patient cases	Case-mix adjusted number of cases serve as a
Hospitals (primary)	Y_2 No. patient cases	severity adjusted output for specialized and
Hospitals (psychiatric)	Y_3 No. patient days	primary hospitals. The output of
Rehab clinic	Y_4 No. patient days	rehabilitation and psychiatric clinics, nursing
Nursing homes	Y_5 No. patient days	homes and retirement homes is measured by
Retirement homes	Y_6 No. patient days	the number of patient days.
(5) Transportation		
Cantonal roads	Y_1 Road length (km)	Road length serves as a proxy for
Communal roads	Y_2 Road length (km)	maintenance. Number of cars is used as a
Road utilization	Y_3 No. cars	utilization indicator for the roads.
(6) Public economy		
Farming	Y_1 Farming area (km2)	The assumption is that farming and Forest
	Y_2 Mountain area (km2)	areas serve recreation purposes. The share of
	Y_3 Organic area (km2)	mountain area and organic farming area
Forestry	Y_4 Forest area (km2)	adjusts for differences in quality.

canton c, $(c = 1, ..., 26)$ in year t, $(t = 2000, ..., 2004)$. They are related to the output indicators $(Y^j_{k,ct}, k = 1, ..., K_j$ for $2 \leq K_j \leq 6)$ and a time trend $(trend^j_t, t = 1, ..., 5)$,

$$
\begin{aligned}
C^1_{ct} &= \beta^1_0 + \beta^1_1 Y^1_{1,ct} + \beta^1_2 Y^1_{2,ct} + \cdots + \beta^1_{K_1} Y^1_{K_1,ct} + \alpha^1 trend^1_t + \varepsilon^1_{ct} \\
&\vdots \\
C^6_{ct} &= \beta^6_0 + \beta^6_1 Y^6_{1,ct} + \beta^6_2 Y^6_{2,ct} + \cdots + \beta^6_{K_6} Y^6_{K_6,ct} + \alpha^6 trend^6_t + \varepsilon^6_{ct}. \qquad (2.5)
\end{aligned}
$$

Since cantons are exposed to similar shocks, error terms ε^j_{it} are likely to be correlated, calling for SURE (Seemingly Unrelated Regression Estimation).

Depending on the service category, SURE confirms the relevance of the selected 22 output indicators and the correlation between the service categories. Pertinent econometric results are shown in Table 2.3 together with the correlation matrix of the residuals. Two criteria were applied to judge the relevance of the output indicators. First, they need to be positively related to cost as a summary measure of input qualities, in keeping with the production theory laid out in Section 2.3. Second, they should importantly contribute to the explanatory power of the cost driver analysis, indicated by the significance level of their coefficients.

Table 2.3: Seemingly Unrelated Regression Results for Six Public Service Categories[a)]

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Y_1	9.74E+04***	-7.18E+04**	1.25E+08***	1.61E+05***	3.17E+04***	2.50E+00***
	2.04E+04	2.22E+04	1.91E+07	2.42E+04	4.83E+03	9.33E+02
Y_2	1.55E+05***	1.76E+05***	2.19E+08***	3.07E+05***	2.17E+04***	7.80E+03***
	2.14E+04	5.19E+04	1.35E+07	2.26E+04	5.85E+03	1.82E+03
Y_3	-	1.69E+05**	5.79E+07***	4.78E+04***	6.17E-01***	1.29E+04***
		5.49E+04	1.28E+07	9.75E+03	3.10E+03	2.56E+03
Y_4	-	-	7.06E+07***	3.98E+03	-	1.67E+04***
			1.16E+07	1.08E+04		1.88E+03
Y_5	-	-	-	1.08E+05**	-	
				3.93E+04		
Y_6	-	-	-	2.38E+04	-	-
				1.72E+04		
adj. R^2	0.94	0.91	0.98	0.96	0.93	0.99
Obs.	130	130	130	130	130	130
Correlation Matrix of the Residuals						
(1)	1.00	0.20	0.03	0.37	-0.17	0.03
(2)		1.00	0.81	0.46	0.48	0.09
(3)			1.00	0.41	0.53	0.05
(4)				1.00	0.09	-0.15
(5)					1.00	0.17
(6)						1.00

(1) Administration, (2) Public safety, (3) Education, (4) Health, (5) Transportation, (6) Public economy.
[a)] standard errors in parentheses, time dummies not shown.
,* Significant at the 95% and 99% confidence level, respectively.

With one exception, the output indicators are positively related to cost. The negative sign of β_1^2 in Table 2.3 (number of *delinquencies*) could be the result of systematic measurement error. Some cantons report the many petty of cases (which cause little expense), while others limit their reporting to the relatively few major offenses (which cost a lot). These differences may induce a negative partial correlation between expenditure on public safety and the number of delinquencies. Furthermore, except for the two coefficients β_4^4 (*rehab clinics*) and β_6^4 (*retirement homes*), all indicators are significant at the 95 percent confidence level. Both indicators are nevertheless retained for reasons of completeness. Finally, the adjusted R^2 reaches at least 90 percent, confirming the combined relevance of the selected indicators.

The Output Performance Index

Including all 22 indicators is still not possible in an annual DEA with 26 cantons. This is because DEA necessarily identifies at least one canton as efficient w.r.t. the 22 output-cost combinations. Thus, at least 22 out of 26 cantons would be identified as efficient, reducing the power of the analysis. One possibility is to aggregate the retained 22 output indicators $Y_{k,ct}^j$ to a performance index Ψ_{ct}^j for each service category ($j = 1, .., 6$), resulting in

$$\Psi_{ct}^j = \sum_{k=1}^{K_j} Y_{k,ct}^j * p_k^j. \tag{2.6}$$

In this calculation, the problematic output indicator *delinquencies* is subjected to a linear monotone transformation such that $\widehat{Y_{1,ct}^2} = -Y_{1,ct}^2 + r \geq 0$ (see Seiford and Zhu, 2002). As to the weighting parameters p_k^j pertaining to the two to six output indicators per service category, there are two alternatives. One is to use the estimated β_k^j from eqs. (2.5). Here, all cantons get the same weighting parameters p_k^j. A more flexible variant is to use the canton-specific shadow prices υ_c from a DEA applied to each of the six service categories (see Section 2.3 again). These shadow prices reflect the marginal cost of expanding a particular service by one unit produced by an efficient DMU. For the inefficient DMUs, a radial projection onto the efficiency frontier permits to determine the pertinent shadow prices. In this variant, no canton is discriminated because of an inappropriate weighting parameter p_k^j would be the case in the first alternative. For each canton, the aggregation is based on those cantonal specific weighting parameters υ_c, which maximize their output performance index. For reasons of consistency, this alternative is retained for our analysis. The results of this calculation (with Ψ_c^j values for 2004) are displayed in Table 2.4.

Table 2.4: Output Performance Indicators Ψ_i^j, 26 Swiss Cantons (2004)

Cantons	(1)	(2)	(3)	(4)	(5)	(6)
ZH	804,351	609,182	1,438,309	2,414,338	293,337	182,691
BE	608,731	467,307	1,018,589	2,279,872	270,257	535,600
LU	226,021	160,632	534,219	490,410	84,259	187,310
UR	22,354	16,503	49,531	39,184	12,713	26,513
SZ	87,616	61,522	175,165	128,367	37,009	73,474
OW	21,130	17,194	46,818	33,907	11,161	34,362
NW	28,856	17,994	48,612	35,418	10,613	20,764
GL	25,754	19,353	51,962	61,451	10,999	29,547
ZG	128,873	47,613	134,137	132,750	27,843	31,162
FR	160,165	113,271	363,600	343,419	75,380	180,878
SO	157,621	112,625	312,784	269,547	65,085	85,709
BS	123,028	104,393	232,726	634,082	31,674	1,302
BL	169,043	120,690	383,400	317,776	62,031	58,940
SH	47,015	34,727	96,062	113,555	31,302	36,222
AR	33,668	24,723	74,755	99,397	17,765	43,370
AI	10,082	7,635	29,755	15,113	4,366	24,714
SG	292,344	208,803	620,278	668,789	108,448	205,445
GR	120,802	130,689	221,249	300,494	97,956	306,163
AG	360,075	255,663	858,094	691,775	145,986	146,288
TG	148,445	105,537	332,003	360,721	71,424	117,408
TI	220,732	185,993	420,807	654,548	100,196	84,566
VD	413,159	327,087	986,906	1,411,360	188,657	274,138
VS	183,488	174,242	399,383	459,095	134,423	138,802
NE	106,986	80,628	270,524	284,766	44,079	91,707
GE	288,564	204,038	635,494	1,508,138	106,019	24,811
JU	44,022	31,404	95,142	121,188	29,632	106,580

(1) Administration, (2) Public safety, (3) Education, (4) Health, (5) Transportation, (6) Public economy.

While the numbers are difficult to interpret in general, the entries for administration (col. 1) reflect size of the cantonal population served because the two output indicators are population and number of firms.

2.4.3 Input Variables

The inputs are measured as real expenditure (CHF of 2000) on the six service categories. This is a widespread practice (see Afonso et al., 2006 and De Borger and Kerstens, 1996). For the categories *transportation* and *health*, only operating expenses are included (total expenditure minus investments in new infrastructure) because annual investments contain a strong transitory component.

2.4.4 Determinants of DEA Efficiency

Recall the two research questions,

(1) Does a fiscal equalization program as sizable as the Switzerland contain incentives to provide public goods less efficiently, creating a trade-off between equity and efficiency?

(2) Does it matter whether transfer payments are earmarked or not?

The first question is investigated using three models. Model (A) relates DEA efficiency scores to the financial potential, which determines the amount of fiscal equalization between cantons. Model (B) checks whether this influence depends only on the size of the financial flows regardless of their direction. In model (C), fiscal equalization paid and received is allowed to have an asymmetric impact on efficiency. The explanatory variables are defined as follows (endogeneity issues are addressed in Section 2.5.2 below).

- *Index of financial potential* ($F.POT$): The Swiss fiscal equalization program is based on this indicator, with higher value implying less federal financial assistance (see Section 2.2). It is used in model (A).

- *Index of financial equalization* ($F.EQ$): $F.EQ$ is a modification of $F.POT$. It measures the absolute value of the deviation from the value α at which no aid is contributed or received; formally, $F.EQ = abs[F.POT - \alpha]$. Note that α differs from the mean value of $F.POT$. The higher this index value, the larger is the amount of fiscal equalization. It is used in model (B).

- *Dummies for paying and receiving cantons* ($F.GIV = 1$, $F.REC = 1$): $F.GIV$ equals 1 for cantons who are payers, while $F.REC$ equals 1 for those who are recipients. Cantons which are neither recipients nor payers constitute the benchmark group in both cases. These variables appear in model (C).

The second research question calls for the introduction of

- *Subsidies per capita* ($SUBS$): This variable measures earmarked payments, which are suspected to induce a particularly high degree of inefficiency (see Section 2.2 again).

In addition, the following variables serve to control for other influences on cantonal efficiency scores that cannot be controlled for in the DEA but could influence efficiency scores.

- *Direct democracy* ($DIR.DEM$): The degree of direct democratic control (popular initiatives, mandatory referenda on expensive public projects) was already found to be relevant by Pommerehne and Zweifel (1991) in the context of tax evasion. More recently, Fischer (2004) and Feld and Matsusaka (2003) found the amount of public services provided to be negatively related to an index of democratic control developed by Stutzer (1999). This index is used here as well, with the expectation of a positive relationship with efficiency.

- *Decentralization* (DEC): Decentralized provision of public services within a canton has an ambiguous effect on efficiency. On the one hand, it might cause a lack of human and technical resources in small cantons, resulting in higher cost of administration (see e.g. Smith, 1985). On the other hand, Tiebout (1956) argues that decentralization facilitates competition, which fosters efficiency. In this work, DEC is the share of cantonal expenditure that is transferred to the communities.

- *Income per capita* ($INCOME$): This is a component of $F.POT$ that according to De Borger and Kerstens (1996) has additional information content. They predict that efficiency of local government decreases with increasing income per capita because citizens in high-wage jurisdictions face high opportunity costs when trying to monitor the efficiency of public good provision.

- *Tax burden* (TAX): This component of $F.POT$ has additional information content as well. In line with Tiebout (1956), a canton's efficiency awareness is predicted to increase with a stronger participation in tax competition. Since a low value of TAX indicates a strong engagement in tax competition, it is hypothesized to go along with a high degree of efficiency, ceteris paribus.

- *Disparities* ($TOPOGR$, $I.STRUCT$, and $POP.STRUCT$): These variables reflect exogenously given disparities, which are expected to cause higher cost and hence lower efficiency in the provision of public services. They enter the new fiscal equalization formula. $TOPOGR$ adjusts for geographic differences while $I.STRUCT$ controls for difference of community size, the employment rate, and population density. $POP.STRUCT$ denotes the shares of immigrants and citizens older than 80 years, with equal weights.

- *Cost of housing* ($P.HOUS$): The cost of housing differs substantially between cantons. It is an important component of the cost of living, which is adjusted for in the wages of public employees and hence influences the cost of providing public services.

- *Culture* ($CULT.F = 1$): The French- and German-speaking parts of Switzerland differ in many ways, possibly also in terms of efficiency (Fischer, 2004). Thus, $CULT.F = 1$ if the canton is predominantly French-speaking.

- *Year of observation* ($Y\_2001 = 1$, $Y\_2002 = 1$, $Y\_2003 = 1$, and $Y\_2004 = 1$): This set of dummy variables indicates the year of observation (base year is 2000).

2.5 Empirical Results

This section first discusses the robust DEA efficiency scores. The assumption (to be relaxed below) is that the 26 cantons belong to the same universe, meaning that all cantons face the same circumstances in their provision of public goods. In a second step, efficiency scores are related to fiscal equalization and other socioeconomic factors of interest.

2.5.1 DEA Analysis

With the six output indicators derived from eq. (2.6) and expenditures changing from year to year, an annual DEA for the years 2000 to 2004 can be performed. Table 2.5 shows the results for the year 2004. The robust efficiency scores are calculated under the assumption of constant returns to scale, indicating potential cost improvements achievable by a radial movement to a technically and scale-efficient reference point on the frontier. There are two super-efficient cases that are assigned a score of 1.00 (see Section 2.3 again).

Table 2.5: DEA Efficiency Scores, 26 Swiss Cantons (2004)

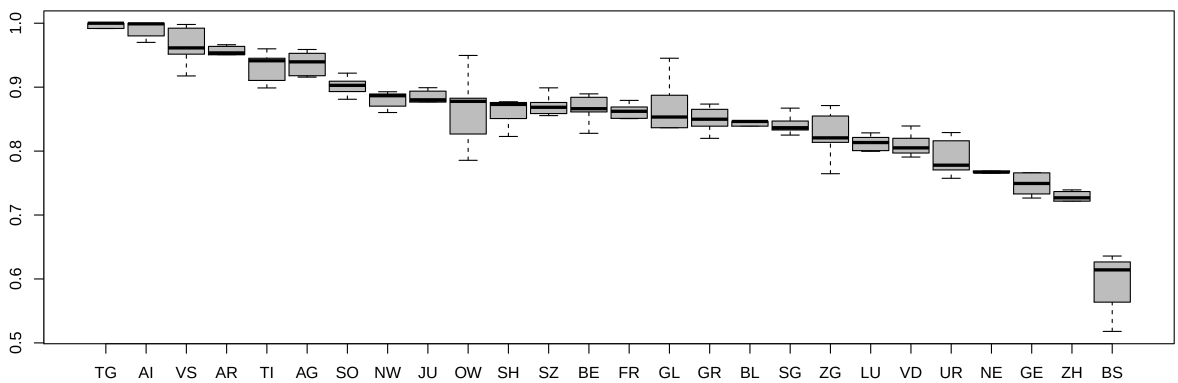
Cantons	Rank	(1-6)$^{a)}$	(1)	(2)	(3)	(4)	(5)	(6)	SD
ZH	25	0.74	0.71	0.42	0.67	0.76	0.53	0.84	0.15
BE	11	0.88	0.87	0.70	0.76	0.81	0.71	0.88	0.08
LU	20	0.82	0.64	0.74	0.91	0.63	0.67	0.79	0.11
UR	22	0.82	0.71	0.65	0.82	0.76	0.69	0.73	0.06
SZ	15	0.86	[1.00]	0.75	0.81	0.56	0.71	0.75	0.14
OW	5	0.95	0.78	0.97	0.93	0.72	0.88	0.78	0.10
NW	10	0.89	[1.00]	0.79	0.78	0.64	0.78	0.77	0.12
GL	6	0.95	0.81	0.91	0.85	0.97	0.74	0.77	0.09
ZG	16	0.85	0.95	0.51	0.69	0.81	0.69	0.92	0.16
FR	14	0.86	0.79	0.74	0.93	0.69	0.67	0.77	0.09
SO	12	0.88	0.78	0.73	0.81	0.61	0.81	0.96	0.11
BS	26	0.64	0.84	0.42	0.78	0.72	0.43	0.21	0.25
BL	18	0.85	0.68	0.70	0.84	0.69	0.62	[1.00]	0.14
SH	13	0.88	0.71	0.59	0.84	0.73	0.99	0.82	0.13
AR	8	0.93	0.64	0.91	0.86	0.93	0.61	0.98	0.16
AI	2	0.97	0.92	0.86	0.98	0.83	0.61	0.97	0.14
SG	19	0.84	0.89	0.65	0.80	0.74	0.49	0.89	0.16
GR	17	0.85	0.61	0.90	0.88	0.63	0.52	[1.00]	0.19
AG	3	0.95	0.82	0.82	[1.00]	0.59	0.95	0.91	0.14
TG	[1]	[1.00]	0.84	0.66	0.85	[1.00]	[1.00]	0.99	0.14
TI	7	0.95	0.62	0.79	[1.00]	0.85	0.78	[1.00]	0.14
VD	21	0.82	0.67	0.60	[1.00]	0.55	0.71	0.84	0.16
VS	4	0.95	0.82	[1.00]	0.96	0.76	0.77	0.78	0.10
NE	23	0.77	0.64	0.55	0.98	0.39	0.53	[1.00]	0.25
GE	24	0.75	0.42	0.38	0.91	[1.00]	0.65	0.64	0.25
JU	9	0.89	0.70	0.81	0.88	0.45	0.93	[1.00]	0.20
No. Eff.		1	2	1	3	2	1	5	
Outliers		0	0	0	0	TG	0	TI	
Mean		0.87	0.76	0.71	0.87	0.72	0.71	0.85	
Min		0.64	0.42	0.38	0.67	0.39	0.43	0.21	
SD		0.08	0.14	0.17	0.09	0.16	0.15	0.17	

(1)Administration, (2) Public safety, (3) Education, (4) Health, (5) Transportation, (6) Public economy.
$^{a)}$ Mean of the six categories, normalized by the maximum value.

Starting with the overall scores, the rural canton of Thurgovia (TG) attains 100 percent technical efficiency (score of 1.00). Two more cantons (again rural) come close, viz. Appenzell Inner-Rhodes (AI, 0.97) and Argovia (AG, 0.97). Indeed, 30 percent of all cantons have a performance score higher than 0.90. At the other extreme, Basel-City (BS) is identified as the most inefficient canton (0.64). Thus, its expenditure could have been lowered by 37 percent while still maintaining the same output level. Other urban cantons, viz. Zurich (ZH, 0.74), Geneva (GE, 0.75) already perform much better. However, differences between rural and urban cantons are not surprising. The well-known disparities caused by higher population densities and more complex industry structures, which by the way are taken into account in the fiscal equalization program, cannot be incorporated in DEA. But the second step analysis adjusts for it with three variables from the new fiscal equalization program to enable unbiased estimates of the hypotheses.

The question arises of whether the year 2004 is representative of the observation period 2000 to 2004. Figure 2.2 provides an answer, ranking cantons according to their five-year median values along with their estimated quartile ranges and 95 percent confidence bands. The findings of Table 2.5 are confirmed in that TG remains leader while BS consistently

Figure 2.2: Overall Efficiency Scores, 26 Swiss Cantons (2000-4)[a)]



[a)]For the acronyms, see footnote No. 1.

is last. While changes in ranking do occur (see the overlapping interquartile ranges), they are very limited. One reason for volatility over time could be investment in infrastructure. For example, ZG shows an improvement from rank 19 in 2000 to 15 in 2004 but drops to place 23 in 2003, because of spending heavily on investment without charging projects to the capital account. Yet, comparable GL with a similar degree of volatility in performance achieved a consistent improvement over the five years [from 0.84 (rank 18) to 0.94 (rank 6)]. In sum, variations over time are too limited and unsystematic to undermine the robustness of the overall ranking.

Another question of interest is whether the leader TG is the champion in all six categories of public service distinguished. If this were the case, Tiebout competition would unfold with full vigor. However, Table 2.5 shows that TG has a low efficiency score in public safety (0.66). Conversely, last-ranked BS does attain an average value in administration (0.84), permitting cantonal government to cater to voters especially interested in administrative services. Moreover, low overall scores do not necessarily go along with high standard errors across the six categories (see last column of Table 2.5). Bottom-ranked BS has a high standard error of 0.25 while UR with rank 23 has one of only 0.06. Thus, small and homogenous UR can survive Tiebout competition since neighboring (more urban) LU has twice as much variation [SD 0.11], while its rank is almost the same. In sum, Tiebout competition is limited even in a country as markedly federalist as Switzerland.

In a federal state, another major issue is centralization vs. decentralization. In the case of Switzerland, the debate has been focusing on education (see Barankay and Lockwood, 2007). Lack of coordination between the cantons has been cited as a reason for the rather mediocre performance of the Swiss educational system in the PISA study (OECD, 2006). However, these criticisms might be overstated. The average performance score for education (3) is 0.87 (SD 0.09). This beats the score of 0.76 (SD 0.14) for public administration (1), which is generally believed to perform well in international comparison.

2.5.2 Estimation of the Determinants of DEA Efficiency

Next, it is of interest to see whether fiscal equalization has an influence on the efficiency scores of the 26 cantons over the years 2000 to 2004. In total, 130 observations (26*5) are available for estimating eq. (2.4) of Section 2.3. Disparities in the provision of public goods are reflected by the indicators discussed in Section 2.4.4.

Estimation results for the three models outlined in Section 2.4.4 are displayed in Table 2.6, after performing tests for endogeneity, heteroscedasticity, and nonlinearity. Fiscal equalization could be endogenous to efficiency because highly efficient jurisdictions are made to contribute to the program. However, a Hausman test does not suggest rejection of the exogeneity assumption. This is not really surprising because the Swiss fiscal equalization is not adjusted every year, possibly making an observation period of five years too short for detecting reverse causality. Heteroscedasticity is not a problem either according to a Breusch-Pagan test. Finally, linearity need not be rejected with the exception of $SUBS2$, $TAX2$ and $POP.STRUCT2$. *Earmarked payments* as well as *tax burden*, *decentralization*, and *population structure* have a nonlinear influence on the performance of the cantons. Several interaction terms proved significant, too; their inclusion does not markedly affect parameter estimates, however. Estimation results turn out to be robust for the three models. Most of the variables have expected signs and are significant at the 90 percent confidence level or better.

Table 2.6: Tobit Estimates of DEA Efficiency Scores

	Model (A)[a]		Model (B)[a]		Model (C)[a]	
Variables	Coef	Elasticity[b]	Coef	Elasticity[b]	Coef	Elasticity[b]
F.POT	-5.8E-03***	-7.0E-01				
F.EQ			-7.6E-04***	-3.8E-02		
F.GIV					-2.1E-01*	-8.4E-02
F.REC					-2.8E-02	-2.4E-02
SUBS	-1.6E-04***	-2.7E-01	-1.8E-04***	-3.0E-01	-1.3E-04***	-2.2E-01
DIR.DEM	-3.0E-01***	-1.7E+00	-7.1E-02***	-3.9E-01	-1.4E-01***	-7.9E-01
DEC	2.0E+00**	1.1E+00	3.3E+00***	1.7E+00	2.9E+00**	1.5E+00
INCOME	-2.8E-04	-3.7E-02				
TAX	-5.6E-03***	-7.9E-01	-6.4E-03***	-9.0E-01	-7.7E-03***	-1.1E+00
TOPOGR	-1.4E-04**	-2.4E-02	-6.2E-06	-1.1E-03	-5.7E-05	-1.0E-02
I.STRUCT	-3.3E-01***	-7.2E-01	-2.3E-01***	-5.0E-01	-2.3E-01***	-4.8E-01
POP.STRUCT	-3.1E-01***	-5.3E-01	-2.7E-01***	-4.8E-01	-2.6E-01***	-4.6E-01
P.HOUS	-6.3E-01***	-8.5E-01	-6.2E-01***	-8.4E-01	-6.5E-01***	-8.9E-01
CULT.F	-6.0E-02***	-1.8E-02	-4.1E-02***	-1.2E-02	-4.0E-02***	-1.2E-02
*SUBS*2	2.3E-08***	8.8E-02	1.5E-08***	5.6E-02	1.3E-08***	5.0E-02
*TAX*2	3.1E-05***	4.8E-01	3.2E-05***	4.9E-01	3.9E-05***	6.0E-01
*POP.STRUCT*2	7.9E-02***	2.7E-01	6.3E-02***	2.1E-01	5.8E-02***	2.0E-01
*DEC*2	-5.0E+00***	-1.1E+00	-5.0E+00***	-1.1E+00	-5.4E+00***	-1.2E+00
SUBS : *F.POT*	1.2E-07	1.4E-02				
SUBS : *F.EQ*			1.1E-06***	7.7E-02		
SUBS : *F.GIV*					2.1E-05	5.9E-03
SUBS : *F.REC*					2.3E-05	3.0E-02
F.POT : *DIR.DEM*	1.3E-03***	6.4E-01				
F.GIV : *DIR.DEM*					4.0E-02**	6.9E-02
F.REC : *DIR.DEM*					2.8E-05	9.3E-05
DIR.DEM : *DEC*	4.2E-01***	8.9E-01	1.2E-01***	2.6E-01	2.9E-01***	6.2E-01
I.STRUCT : *P.HOUS*	2.9E-01***	6.5E-01	2.0E-01***	4.5E-01	2.0E-01***	4.5E-01
Observation	114		114		114	
L-Likelihood	384.1		374.9		372.2	

*,**,*** Significant at the 90%, 95% and 99% confidence level, respectively.
[a] Time dummies for the years 2004, 2003, 2002 and 2001 are not shown.
[b] Elasticities evaluated at sample means.

In model (A), a negative sign is obtained for $F.POT$. Use of index of financial potential that determines fiscal equalization payments therefore seems to lower efficiency systematically (elasticity -0.7) after controlling for exogenously given disparities and other variables affecting the cost of public good provision. Thus, cantons with high financial potential may have an incentive to underperform. In model (B), the absolute value of payments enters with $F.EQ$. Not surprisingly, $F.EQ$ has a significantly negative sign too, suggesting that fiscal equalization as such lowers technical efficiency in the provision of public goods. Finally, model (C) indicates that paying cantons (elasticity -0.084) are more influenced than receiving cantons (elasticity -0.024) with regard to efficiency.

In sum, the evidence of Table 2.6 provides an answer to question (1) of Section 2.2 by supporting the notion that fiscal equalization undermines cantonal efficiency in Switzerland for both receivers and payers, but even more for payers, who are the cantons with high financial potential. This difference is intuitive because payers have more reason to respond to fiscal equalization with inefficiency than receivers. Expecting no benefit from redistribution, they rather waste their money than to give it to financially disadvantaged cantons. Thus, any public good with a positive net benefit is provided, whereas only those with above-average net benefits contribute to the canton's technical efficiency. Being financially constrained, receiving cantons want to ensure that their most productive projects are financed; they extend this list only in order to justify their need for redistribution. While estimated elasticities are below one throughout, fiscal equalization in the case of Switzerland does give rise to the equity-efficiency trade-off described Stiglitz (1988).

Question (2) of Section 2.2 asks whether earmarked federal subsidies ($SUBS$) have an especially strong (negative) effect on cantonal efficiency. Whereas general payments can be used by the canton where it believes to generate the highest benefit for its citizens, earmarked subsidies may result in gold plating of projects and hence inefficiency. Indeed, Table 2.6 shows $SUBS$ to have a negative sign in all three models with estimated elasticities between -0.2 and -0.3. Therefore, subsidies may encourage inefficiency, as claimed in the Swiss case e.g. by Frey et al. (1994). Therefore, the new equalization formula of 2008 which minimizes earmarked payments has the potential to reduce technical inefficiency in the provision of public goods compared to its predecessor.

Some of the other explanatory variables are of interest as well. Foremost, $DIR.DEM$ and DEC, which capture two unique features of Switzerland, contradict theoretical expectations. The negative sign of $DIR.DEM$ suggests that direct democratic control lowers rather than increases technical efficiency. This seems to contradict the findings of Fischer (2004) as well as Feld and Matsusaka (2003), who however studied the amount of public services provided rather than technical efficiency. Still, lower amounts can go along with lower efficiency if direct democracy should mainly delay (notably through referenda) planning that is "on target" in terms of efficiency. On the other hand, decentralization has the expected effect in that the coefficient of DEC is positive throughout, confirming Barankay and Lockwood (2007) who examined the impact of decentralization on productive efficiency in public education. The negative effects emphasized by Smith (1985) apparently are more than compensated by the positive ones due to Tiebout competition, which however are subject to diminishing marginal returns (see the negative coefficient of $DEC2$).

In addition, TAX shows the expected negative sign, suggesting that cantons with a low tax burden exhibit higher performance, a state of affairs conducive to strong Tiebout competition. However, the positive sign of the quadratic term points to a rapidly diminishing effect as soon as the tax burden starts to increase, with the critical value of 90.32 in model (A) and 100 in model (B), respectively (the average tax burden of Switzerland is set to 100). The positive sign of TAX found by De Borger and Kerstens (1996) therefore also holds for Switzerland as soon as it exceeds the average. Thus, both extremely low and high tax burdens cause efficiency gains, because of tax competition on the one hand and because of increasing monitoring by citizens on the other.

Finally, it is of interest for policy to know whether the determinants entering the new fiscal equalization formula ($TOPOGR$, $I.STRUCT$, and $POP.STRUCT$) to adjust for resource disparities are relevant or not. The three variables are negatively related to DEA efficiency scores regardless of model specification. Therefore, the 2008 reform is likely to achieve its objective because it introduces exogenous factors in the equalization formula that seem to have a significant influence on the heterogeneity of public good provision. Finally, the negative coefficient of $P.HOUS$ shows that the cost of housing factors into the cost of public services and hence inefficiency. Since it is largely exogenous, it could also be included in the fiscal equalization formula.

2.6 Concluding Remarks

The purpose of this paper was to measure efficiency in the provision of public services applying Data Envelopment Analysis (DEA), which maximizes the distance between an output bundle and an input bundle. The country analyzed is Switzerland, which is characterized by a high degree of federalism permitting Tiebout competition on the one hand and a sizable fiscal equalization program on the other. DEA shadow prices serve to derive the weights for aggregating the six public service categories into an overall output indicator for the 26 cantons, while inputs are equated to their real expenditure over the years 2000 to 2004. In a second step, DEA efficiency scores are related to the indicator 'financial potential' which governs the Swiss fiscal equalization scheme designed to alleviate disparities between cantons.

The main results are the following. First, efficiency scores indicate better performance of small rural cantons than of urban ones and are robust over the five years investigated. A comparison over the six service categories further shows that cantons with a high overall performance do not automatically outperform in all of them, preventing any one of them from becoming dominant in Tiebout competition. Second, financial equalization is negatively related to cantonal efficiency, with an especially marked effect on payers. Schemes designed to mitigate disparities that are deemed unacceptable not only by politicians but the citizenry as well (the pertinent constitutional amendment survived a popular referendum in the case of Switzerland) may thus have the undesirable side effect of undermining incentives for efficiency. Jurisdictions who are payers and receivers both seek to keep their 'financial potential' low – the former because this serves to ease their burden, the latter because they expect to receive more transfer payments and subsidies notably by producing public services at higher than minimum cost. Therefore, the equity-efficiency trade-off noted by Stiglitz (1988) seems indeed to exist in the case of Switzerland. Third, earmarked federal subsidies (the main component of transfer payments prior to the 2008 reform) are negatively related to cantonal efficiency as well. Substitution of these earmarked payments by freely disposable lump-sum ones as part of the new equalization program implemented in 2008 is therefore likely to enhance cantonal performance.

This analysis suffers from several limitations. Above all, DEA efficiency scores constitute a technocratic measure, being silent on the question of whether the services provided reflect the preferences of citizens. Also, some of the explanatory variables used to predict efficiency scores may not be fully exogenous in the long term. In particular, $INCOME$ possibly not only influences efficiency as a taste variable but could be the consequence of cantonal efficiency as well. In spite of these limitations, the analysis not only identifies the equity-efficiency trade-off that federally organized countries (such as Switzerland) face when implementing a fiscal equalization scheme but also provides guidance on how to structure it in terms of earmarked and freely disposable payments.

CHAPTER III

Accounting for Heterogeneity in the Measurement of Hospital Performance

Philippe K. Widmer, Peter Zweifel, and Mehdi Farsi

Abstract: With prospective payment of hospitals becoming more common, measuring their performance is gaining in importance. However, the standard cost frontier model yields biased efficiency scores because it ignores technological heterogeneity between hospitals. In this paper, efficiency scores are derived from a random intercept and an extended random parameter frontier model, designed to overcome the problem of unobserved heterogeneity in stochastic frontier analysis. Using a sample of 100 Swiss hospitals covering the years 2004 to 2007 and applying Bayesian inference, significant heterogeneity is found, suggesting rejection of the standard cost frontier model. Estimated inefficiency decreases even below the 14 percent reported by Hollingsworth (2008) for European countries. Accounting for unobserved heterogeneity would make hospitals rated below 85 percent efficiency according to the standard model gain up to 12 percentage points, serving to highlight the importance of heterogeneity correction in the estimation of hospital performance.

Keywords: Hospital efficiency, unobserved heterogeneity, Bayesian inference, Switzerland, Stochastic Frontier Analysis

JEL classification: C11; C33; D24; I18

3

Accounting for Heterogeneity in the Measurement of Hospital Performance

3.1 Introduction

Performance-based prospective hospital payment has recently been introduced in several countries. It has greatly increased the importance of accurately measuring firm-specific performance, defined here in terms of operating costs per casemix-adjusted patient case. The challenge for policy makers is to pay for efficiency; however, this calls for filtering out differences that are caused by inefficiencies rather than heterogeneities due to exogenous influences.

In response to this need, there has been a growing body of research into the determinants of cost variability between hospitals (see Hollingsworth, 2008, Jacobs et al., 2006, and Worthington, 2004 for overviews of the literature). Specifically, the meta-analysis by Hollingsworth (2008) finds evidence of inefficiency in the hospitals of the United States and several European countries, amounting to a potential for cost reduction of 18 percent and 14 percent, respectively. Compared to these estimates, Switzerland is on the high side with Steinmann and Zweifel (2003), based on a Data Envelopment Analysis (DEA), coming up with 30 percent. In a comparison with Germany (the land of Saxony), Steinmann et al. (2004) once more found Swiss hospitals to be relatively inefficient. Using a Stochastic Frontier Analysis (SFA), Farsi and Filippini (2006) put the potential of cost reduction to 20 percent on average, which would translate into 7 percent of Switzerland's total health care expenditure (55 bn CHF) in 2007. However, these estimates do not account for heterogeneity in production technology, which is likely to be particularly marked due to Swiss federalism.

Efficiency scores from the articles cited above are simply defined as the ratio of observed cost to a value on the estimated single technology cost frontier (Farrell, 1957).[1] In the case of SFA introduced by Meeusen and van den Broeck (1977) and Aigner et al. (1977), this ratio is given by

$$\frac{C_{it}}{C_S(X_{it};\alpha,\beta)} = e^{u_i+v_{it}}, \tag{3.1}$$

with C_{it} is the (arithmetic) cost of hospital $i = 1, ..., N$ at time $t = 1, ..., T$, $C_S(X_{it};\alpha,\beta)$ is an estimated minimum cost for outputs and input prices comprised in a $NT \times (K+1)$ matrix X, α is the unknown intercept, β is a $K \times 1$ vector of unknown slope parameters of the cost function, u_i is a random term with positive values only reflecting inefficiency, and v_{it} a conventional random error (see Section 3.2 below for details). Any difference in technology is captured in the composite error term $u_i + v_{it}$, which could bias the estimates of the inefficiency scores u_i (see also Greene, 2004a). This is particularly the case for Switzerland, where hospitals have to operate in different regulatory environments, causing them to provide health care services using different technologies. Controlling for heterogeneity may therefore lead to inefficiency estimates that are more in line with those of the United States and the European countries.

There have been several approaches for dealing with this problem. The first is to introduce fixed effects in the SFM or in the distribution of u_i (see e.g. Worthington, 2004). Since the choice of the dummy variables must rely on observable characteristics of the hospital, this solution is limited to 'observable' heterogeneity, leaving potential for 'unobservable' heterogeneity to bias estimated inefficiency scores. Provided panel data available, a true random effects model can be estimated (see e.g. Farsi et al., 2008, Farsi et al., 2006, Greene, 2005b, and Greene, 2004b). This is a special case of the Random Intercept Frontier Model (RIFM) to be presented below, which enables cost frontiers to vary between hospitals. Still, the RIFM is not without limitations because it only al-

[1] Simple cost ratios, often used for policy purposes, are not sufficient because they neglect both economies of scale and heterogeneity of technology.

lows the intercept α to be stochastic and assumes heterogeneity to be homoscedastic. Additional flexibility is afforded by the Random Parameter Frontier Model (RPFM), a generalization of the SFM and the RIFM, which additionally allows the slopes β vary between hospitals. Implementation of RPFM until recently was hampered by the requirement of large computational power and panel data of sufficient quality. Improvement on both counts have rendered them feasible in the meantime (see e.g. Widmer, 2011, Huang, 2004, Orea and Kumbhakar, 2004, and Tsionas, 2002).

In this paper, we analyze the influence of unobserved heterogeneity between Swiss hospitals using a SFM, a RIFM, and an extended RPFM for SFA. Section 3.2 contains additional details for these models as well as for the Bayesian approach adopted in the model specification and parameter estimation. The database of about 100 Swiss hospitals covering the years 2004 to 2007 is presented in Section 3.3. Estimation results confirm the existence of unobserved heterogeneity in Swiss hospitals, suggesting rejection of the SFM. On average, the SFM overestimates Swiss hospital inefficiency by about 6 percent. Section 3.4 considers the implications of this study for hospital managers and policy makers, and concludes.

3.2 Modeling Unobservable Heterogeneity

It is common practice to define heterogeneity as time-invariant cost variation that is exogenous in the sense that it cannot be manipulated by management at least in the short run. This definition is adopted here.[2] As a benchmark, the specification of the cost function in the presence of observable heterogeneity is presented first (Section 3.2.1); this provides the point of departure for the modeling of unobservable heterogeneity in Section 3.2.2, which likely is the more important component given that the quality of (largely non-profit) hospital management is not easily measured.

[2] On the longer run, the choice of technology can be influenced by hospital management. The presence of inferior technology becomes a component of management inefficiency in this case.

3.2.1 The Standard Frontier Model with Observable Heterogeneity

When heterogeneity is perfectly observable, the SFM can be augmented in a way permitting to estimate consistent efficiency scores (Pitt and Lee, 1981). Let there be repeated observations $t = 1, ..., T$ for all hospitals $i = 1, ..., N$, and let heterogeneity Z_i be completely reflected by a $N \times L$ matrix of observable, time-invariant characteristics. Using a Cobb-Douglas cost function in logs for simplicity, the cost frontier can be specified as

$$C_{it} = \alpha + \gamma' Z_i + \beta' X_{it} + \gamma' X_{it} Z_i + u_i + v_{it}, \quad i = 1, ..., N, \quad t = 1, ..., T, \tag{3.2}$$

$$u_i \overset{iid}{\sim} f_N^+[0, \sigma_u^2] \text{ with } Cov(u_i; \alpha, X_{it}, Z_i, v_{it}) = 0, \tag{3.3}$$

$$v_{it} \overset{iid}{\sim} f_N[0, \sigma_v^2] \text{ with } Cov(v_i; \alpha, X_{it}, Z_i, u_i) = 0, \tag{3.4}$$

with parameters defined in the Introduction section. Note that observable heterogeneity enters the cost function in two ways. First, the intercept α is augmented by a hospital-specific term $\gamma' Z_i$. This is called separable heterogeneity because it captures cost variability that is unrelated to the hospital's technology, such as size of its catchment area.[3] Second, the slopes β pertaining to outputs and factor prices are individualized by $\gamma' X_{it} Z_i$. This is non-separable heterogeneity reflecting differences in technology such as the amount and vintage of capital in use.

In this model, firm-specific inefficiency u_i is assumed to be time-invariant, uncorrelated with the deterministic part of the cost function as well as with random noise v_{it}, and to follow a one-sided distribution supported on the interval $[0, \infty)$ such as the half normal, truncated normal, exponential, gamma or Weibull distribution. Alternatively one could model inefficiency as a time-invariant fixed effect (Schmidt and Sickles, 1984), which allows for correlation with production technology.[4] But because heterogeneity is assumed to be time-invariant here, this variant will not be pursued.

[3] Greene (2005a) permits correlation between separable heterogeneity and production technology in his true fixed effects model.

[4] The fixed effect model has at least two drawbacks: (1) it only measures relative inefficiency, (2) no time-invariant technology parameters are allowed in the cost function.

Given eqs. (3.2) to (3.4) as an option, cost variability in logarithms is given by

$$C_{it} - C_H(X_{it}, Z_i; \alpha, \beta, \gamma) = u_i + v_{it}, \tag{3.5}$$

where $C_H(X_{it}, Z_i; \alpha, \beta, \gamma)$ is individualized for each hospital as shown in eq. (3.2). It enables the hospital regulator to separate cost variability related to differences in technology from inefficiency u_i and random noise v_{it}. The resulting inefficiency scores are no longer biased by heterogeneity and can be used for performance-based reimbursement.

However, this result does not hold as soon as heterogeneity is not completely observable but has an unobservable component Z_i^*, resulting in terms $\gamma' Z_i^*$ and $\gamma' X_{it} Z_i^*$ in the error term of eq. (3.2). If the regulator uses the conventional benchmark C_H of eq. (3.2), the residuals become

$$C_{it} - C_H(X_{it}, Z_i; \alpha, \beta, \gamma) = \gamma' Z_i^* + \gamma' X_{it} Z_i^* + u_i + v_{it}. \tag{3.6}$$

The existence of (time-invariant) unobservable heterogeneity Z_i^* now causes bias in the measurement of hospital performance in two ways:

- Rather than estimating true inefficiency u_i, eq. (3.2) will estimate an artificially augmented inefficiency term $\widetilde{u_{it}} = u_i + \gamma' Z_i^* + \gamma' X_{it} Z_i^*$;
- To the extent that unobserved heterogeneity Z_i^* is correlated with either observed heterogeneity Z_i or outputs and factor prices X_{it} estimates of technology parameters α and β are biased as well since $Cov(\widetilde{u_{it}}, Z_i) \neq 0$, $Cov(\widetilde{u_{it}}, X_{it}) \neq 0$.

With both inefficiency scores and technology parameters biased, prospective payment runs the risk of rewarding some hospitals for being seemingly efficient while punishing others for being seemingly inefficient.

U.S. experience with prospective payment suggests that unobservable heterogeneity could be substantial. If performance-based reimbursement took into account all relevant determinants of hospital cost, one would expect hospitals to discard technologies giving rise to characteristics that are not paid for and to move to the cost-efficient level. However, Keeler (1990) found that U.S. hospitals still have a great deal of unexplained cost variability although prospective payment had been in place since 1983. One reason is that changes in hospital technology are particularly costly, causing hospitals to be slow in adopting new technologies. Evidently, more advanced estimation techniques are necessary to disentangle latent heterogeneity from inefficiency for prospective hospital payment to have the desired efficiency-enhancing effects.[5]

3.2.2 The Random Parameter Model with Unobservable Heterogeneity

The discussion of the preceding subsection led to the conclusion that in the case of hospitals, part of their technological heterogeneity is unobservable for years to come. Therefore, reimbursement arguably should take into account both observable and unobservable heterogeneity. One way to achieve this is the RPFM (the RIFM allows only the intercept rather than all parameters to be random and will not be expounded separately below). A RPFM estimates an individual cost function for each hospital, admitting of both observable and unobservable heterogeneity (see e.g. Widmer, 2011, Greene, 2004b, Huang, 2004, Orea and Kumbhakar, 2004, and Tsionas, 2002). To save on notation, the case of unobservable heterogeneity only is presented below. No ex-ante information on heterogeneity is needed, except for the assumption that it is time-invariant and normally distributed over individual hospitals. This is achieved by introducing a $[(K+1)\times 1]$ vector of time-invariant random variables $w_i \sim N[0, \sigma_w^2]$ that changes the intercept of the cost function to become $\alpha_i = \alpha + w_i$ (separable heterogeneity) and to the slope parameters to become $\beta_i = \beta + w_i$ (non-separable heterogeneity), resulting in

$$C_{it} = (\alpha + w_i) + (\beta + w_i)' X_{it} + u_{it} + v_{it}, \quad \text{or} \tag{3.7}$$

$$C_{it} = \alpha_i + \beta_i' X_{it} + u_{it} + v_{it}. \tag{3.8}$$

In the special case where w_i captures all existing heterogeneity, the RPFM can be transformed back into a SFM by substituting w_i by $\gamma' Z_i$ in eq. (3.7).

5 As stated by Newhouse (1996), one option to overcome this problem is to use non fully prospective reimbursement systems, that are not fully prospective.

However, this specification is somewhat restrictive because it assumes both the intercept and the slopes of the cost function to be time-independent. This neglects the fact that new medical technology affects the whole hospital industry in very much the same way (such as the introduction of SCAT scanners). Denoting these changes by a vector of time dummies M_t, eq. (3.8) can be generalized to read,

$$\begin{aligned} C_{it} &= \alpha_{it} + \beta_{it}' X_{it} + u_{it} + v_{it} \quad \text{with} \\ \alpha_{it} &= \bar{\bar{\alpha}} + \delta' M_t + w_i \quad \text{and} \quad \bar{\alpha}_t = \bar{\bar{\alpha}} + \delta' M_t; \\ \beta_{it} &= \bar{\bar{\beta}} + \delta' M_t + w_i \quad \text{and} \quad \bar{\beta}_t = \bar{\bar{\beta}} + \delta' M_t. \end{aligned} \tag{3.9}$$

This specification allows to disentangle inefficiency from unobservable heterogeneity both variable and time-invariant. Separable heterogeneity is captured by the random intercept $\alpha_{it} = \bar{\bar{\alpha}} + \delta' M_t + w_i$, where $\bar{\bar{\alpha}}$ is the mean intercept over all hospitals. Non-separable heterogeneity in technology parameters is captured by a $(K \times 1)$ vector of hospital-specific parameters $\beta_{it} = \bar{\bar{\beta}} + \delta' M_t + w_i$.

To derive individual effects, assume w_i to follow a multivariate normal distribution

$$w_i \sim f_{MN}\left[\bar{w}, \Sigma\right], \quad \text{with} \;\; \Sigma \sim f_W \begin{bmatrix} \sigma^2_{w_\alpha} & \sigma_{w_\alpha, w_\beta} \\ \sigma_{w_\alpha, w_\beta} & \sigma^2_{w_\beta} \end{bmatrix}, \tag{3.10}$$

with $\bar{w}$ equal to zero. Σ is Wishart distributed with a $[(K+1) \times (K+1)]$ positive definite covariance matrix $S = (\sigma^2_{w_\alpha}, \sigma^2_{w_\beta}, \sigma_{w_\alpha, w_\beta})$, denoting unobserved heterogeneity between hospitals. For $\Sigma = 0$ no time-invariant heterogeneity exists and the random parameter model simplifies to a SFM with no variation in β_{it} and α_{it}.

For the Bayesian estimation to be performed in Section 3.3, the posterior distribution for the random parameter model needs to be derived. It is given by

$$\begin{aligned} p(\alpha,\bar{\bar{\alpha}},\beta,\bar{\bar{\beta}},\delta,u,\Sigma,\sigma_v^{-2},\sigma_u^{-2};C,X,M) \propto\ & p(\bar{\bar{\alpha}},\bar{\bar{\beta}},\delta,\Sigma,\sigma_v^{-2},\sigma_u^{-2}) \\ & \times \prod_{i=1}^{N}\prod_{t=1}^{T}\frac{1}{\sqrt{2\pi\sigma_v^2}}exp\left[-\frac{1}{2\sigma_v^2}\left(C_{it}-[\alpha_{it}+\beta_{it}'X_{it}+u_{it}]\right)^2\right] \\ & \times \prod_{i=1}^{N}(2\pi)^{-K/2}|\Sigma|^{-1/2}exp\left[-\frac{1}{2}\left(w_i-\bar{w}\right)'\Sigma^{-1}\left(w_i-\bar{w}\right)\right] \\ & \times \prod_{i=1}^{N}\prod_{t=1}^{T}p(u_{it},\sigma_u^{-2}), \end{aligned} \tag{3.11}$$

where probability distributions $p(\bar{\bar{\alpha}},\bar{\bar{\beta}},\delta,\Sigma,\sigma_v^{-2},\sigma_u^{-2})$ for the priors remain to be specified in Section 3.3.1. The likelihood function in eq. (3.11) is as in Griffin and Steel (2007), normally distributed with σ_v^2 as the variance of the random noise $v_{it} = C_{it} - [\alpha_{it} + \beta_{it}'X_{it} + u_{it}]$ and with α_{it} and β_{it} as in eq. (3.9). The last term of eq. (3.11) points to a gain in flexibility over classical maximum likelihood applications, where a joint density function of the random noise v and the inefficiency term u is specified. Here, only random noise enters the likelihood function, while inefficiency is estimated along with the other parameters of the cost function.

The random intercept α_{it} and technology parameters β_{it} are estimated at three levels. At the first level, the overall influences on hospital costs ($\bar{\bar{\alpha}}$, $\bar{\bar{\beta}}$) are determined, corresponding to the first factor following the proportionality sign of eq. (3.11); the second-level estimates of time-specific effects ($\bar{\alpha}_t$, $\bar{\beta}_t$) defined in eq. (3.9) and the third-level estimates of individual values (α_{it}, β_{it}) are derived from the multivariate normal distribution shown in eq. (3.11). Finally, inefficiency $p(u,\sigma_u^{-2})$ given by the last factor of eq. (3.11) is estimated at two levels. The first-level estimate corresponds to the population mean. The second-level estimates yield firm-specific inefficiency scores u_{it}.

Note that estimates of the unknown parameters can be derived by the marginal posteriors of eq. (3.11). However, it is not always possible to compute the posteriors analytically. Therefore, iterative Monte Carlo Markov Chain (MCMC) simulation is used, which involves iterative sampling from posterior parameter densities. Here, we use WINBUGS to derive the estimates (see Ntzoufras, 2009 for an introduction to Bayesian analysis with WINBUGS).

The model formulation of eq. (3.9) is quite general, containing many other variants of RPFM cited as special cases. For example, if M_t turn out to be zero for all covariates, it reduces to the one described by Greene (2004b) where all parameters are allowed to vary between hospitals but are constant over time. Furthermore, if additionally only technology parameters contain heterogeneity, it reduces to the one of Tsionas (2002), with a common intercept for all hospitals. If only the intercept controls for heterogeneity, it corresponds to the true random effects model of Greene (2005a) where all cost functions have the same slopes, adjusting only for separable heterogeneity. This is similar to the applied RIFM but with heterogeneity that is constant over time. Finally, if additionally w_i is zero for all parameters, the RPFM reduces to a SFM with no heterogeneity in the cost function.

3.3 Empirical Application to Swiss Hospitals Using Bayesian Inference

In this section, we analyze the effect of unobserved heterogeneity on the performance of Swiss hospitals with three SFA formulations, a Random Parameter Frontier Model (RPFM) outlined in Section 3.2.2, a Random Intercept Frontier Model (RIFM) as a special case of the RPFM, and a Standard Frontier Model (SFM).

3.3.1 Data and Econometric Specification of the Cost Function

The data used in this study are provided by the annual reports of the Federal Office of Public Health. They include 333 Swiss hospitals for the time period 2004 to 2007, comprising information on 5 university hospitals (K111), 23 central hospitals (K112), 27 large regional hospitals (K121), 46 medium regional hospitals (K122), 46 small regional hospitals (K123), 28 specialized surgery hospitals (K231), and sundry hospitals, viz. psychiatric and rehabilitation clinics. In total, 127 of these 333 units are private, non-subsidized hospitals.

In the interest of comparability, the sundry category was discarded. After purging the data from missing values and outliers, an unbalanced panel of 405 observations of sufficient quality is available. Variables are defined as follows:

VC : Variable operational expense per year, in thousands of CHF (VC);

X_1 : CMI-adjusted inpatient cases (PCASES);

X_2 : Revenue from outpatients (OUTP);

X_3 : Price of labor, average wage per employee (PL);

X_4 : No. of beds (BEDS);

S_1 : No. of internship categories (INTERN);

S_2 : No. of specialties (SPEC);

S_3 : Dummy=1 for subsidized public hospitals (SUB);

S_4 : Share of inpatients with supplementary insurance, in percent (INSUR);

M_t : Year dummies, t=2005, 2006, 2007 (base = 2004);

Z_l : Hospital group dummies, l=K111, K112, K121, K122, K231 (base = K123).

Table 3.1: Descriptive Statistics

Variable	Mean	Min	Max	K111	K112	K121	K122	K123	K231
$VC^{1)}$	109,296	3,925	953,586	790,609	190,380	88,122	44,830	19,538	36,387
$PCASES$	7,731	497	52,143	43,046	14,651	7,093	3,796	1,545	2,741
$OUTP^{1)}$	20,499	0	186,174	142,331	42,845	14,753	6,826	2,387	6,698
$PL^{1)}$	101	34	188	106	103	100	100	99	108
$BEDS$	201	12	1,170	893	383	207	108	55	69
$INTERN$	18	0	134	118	32	15	10	5	3
$SPEC$	35	4	86	67	50	36	33	24	18
$SUB^{2)}$	87	0	100	100	100	100	88	80	37
$INSUR^{2)}$	25	3	100	18	19	20	25	27	47

1) in 1,000 CHF, 1 CHF=0.8 USD (2004 exchange rates).
2) in percent, SUB=100 means that 100 percent of all hospitals are subsidized.

Summary statistics are shown in Table 3.1 for all six hospital groups retained. They suggest that technological heterogeneity between Swiss hospital groups indeed influence cost. University hospitals (K111) for example have the highest variable costs ($VC = 790,609$); this can be attributed to their high values of the two major outputs ($PCASES = 43,046$ and $OUTP = 142,331$) and possibly the fact that they are all subsidized ($SUB = 100$ percent) while having a small share of patients with supplementary insurance ($INSUR$ = 18 percent). However, they also have the most internship programs ($INTERN$ = 118) and specialties ($SPEC$ = 67). Specialized hospitals (K231) on the other hand are on average small hospitals with fewer internship programs ($INTERN = 3$) and specialties ($SPEC = 18$) but are mostly non-subsidized ($SUB = 37$ percent) while having a high share of supplementary insured patients ($INSUR = 47$ percent).

With these data, we can specify a cost function where variable cost (VC) depends on two output categories ($PCASES$, $OUTP$), one input price for labor (PL), one capital stock ($BEDS$), and four structural variables ($INTERN$, $SPEC$, SUB, $INSUR$). Although some of them could be interpreted as observable indicators of technology (in particular, the number of specialties offered), they are treated as a category of their own here. Intercept and technology parameters are reflected by a linear function of dummies for different hospital groups (Z) and different time periods (M).[6] The underlying Cobb-Douglas cost function (subscripts $i = 1, ..., N$ and $t = 1, ..., T$ are dropped for simplicity) therefore reads,

$$
\begin{aligned}
lnVC &= \beta_0 + \sum_{m=1}^{4} \beta_m lnX_m + \sum_{z=1}^{3} \beta_z S_z + u + v, \quad \text{with} \\
\beta_{k,\ k=\{0,\ m,\ z\}} &= \bar{\bar{\beta}}_k + \sum_{l=1}^{5} \gamma_{l,k} Z_l + \sum_{\tau=1}^{3} \delta_{\tau,k} M_\tau + w.
\end{aligned}
\tag{3.12}
$$

In order to conduct Bayesian inference from the posterior given in Section 3.2.2, prior distributions need to be specified. The values for the hyperparameters are chosen in a way to imply relatively vague but proper priors. In particular, the priors are assumed to be independent,

$$
p(\bar{\bar{\alpha}}, \bar{\bar{\beta}}, \gamma, \delta, \Sigma, \sigma_v^{-2}, \sigma_u^{-2}) = p(\bar{\bar{\alpha}})p(\bar{\bar{\beta}})p(\gamma)p(\delta)p(\Sigma)p(\sigma_v^{-2})p(\sigma_u^{-2}). \tag{3.13}
$$

Here, $p(\bar{\bar{\alpha}}) = f_N[0, \theta_{\bar{\bar{\alpha}}}]$, $p(\bar{\bar{\beta}}) = f_N[0, \theta_{\bar{\bar{\beta}}}]$, $p(\gamma) = f_N[0, \theta_\gamma]$, and $p(\delta) = f_N[0, \theta_\delta]$ have a normal distribution with mean zero and a diffuse prior for their corresponding precision

[6] Note that average length of stay is not included variable. Its expected value enters the casemix adjustment of $PCASES$. Therefore, deviations from expected value can be interpreted as reflecting management inefficiency u.

θ. The precision of the likelihood function has a gamma distribution $p(\sigma_v^{-2}) = f_G[\mu, \theta_{\sigma_v^{-2}}]$ with diffuse shape and scale parameters. Inefficiency is assumed to be half normally distributed $p(u, \sigma_u^{-2}) = f_N^{+}[0, \sigma_u^{-2}]$ with $\sigma_u^{-2} = f_G[5, (5 * log(\overline{eff})^2)]$. This specification is in line with Griffin and Steel (2007) and Koop et al. (1997), permitting to impose a priori information with regard to mean efficiency, $\overline{eff} = exp(-\overline{u})$. Following the formulation of Griffin and Steel (2007), $\overline{eff} = 0.875$ is assumed for prior efficiency. Finally, the precision of the random parameters is specified as a Wishart distribution $p(\Sigma) = f_W[S]$ in accordance with eq. (3.9) with diffuse prior for the covariance matrix S.

To obtain posterior estimates, MCMC algorithms were run for 100,000 iterations, with the first 50,000 discarded as a burn-in phase. Different assumptions for priors and starting values converged to roughly the same values, suggesting that convergence to the posterior distribution was achieved.

3.3.2 Econometric Results of the Cost Functions and their Cost Variability

Table 3.2 presents estimated means and standard errors of the cost function for the Standard Frontier Model (SFM), the Random Intercept Model (RIFM), and the Random Parameter Model (RPFM). The RPFM is presented with estimates for $\bar{\bar{\alpha}}$ and $\bar{\bar{\beta}}$. Results of second level-estimates of the technology parameters and the intercept are shown in Table 3.5 of the Appendix; they point to a cost shift over time, but without affecting the slope parameters. The three variants in Table 3.2 can be assessed using the DIC information criterion (Spiegelhalter et al., 2002). The lower the DIC-value, the better the goodness of fit of the estimated cost function, indicating that the RPFM has the best model fit, followed by RIFM, with SFM definitely behind.

Nevertheless, the three specifications produce fairly minor variations in technology parameters. Estimates also have the expected sign, with the only exception of BED. Being an indicator of capital stock, it should have a negative sign, which only obtains in the RPFM.

However, the main interest of this research revolves about technological heterogeneity that may not be accounted for in the SFM and its influence on estimated inefficiency

Table 3.2: Econometric Results

	SFM		RIFM		RPFM	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Constant	2.176	(0.279)	3.334	(0.340)	4.948	(1.109)
$PCASES$	0.674	(0.037)	0.554	(0.045)	0.485	(0.115)
$OUTP$	0.025	(0.004)	0.015	(0.004)	0.013	(0.011)
PL	0.254	(0.058)	0.199	(0.049)	0.384	(0.110)
$BEDS$	0.256	(0.039)	0.311	(0.044)	−0.148	(0.127)
$INTERN$	0.004	(0.000)	0.001	(0.001)	0.008	(0.006)
$SPEC$	0.002	(0.001)	0.001	(0.001)	0.001	(0.007)
SUB	0.141	(0.036)	0.191	(0.055)	0.231	(0.529)
$INSUR$	0.004	(0.001)	0.004	(0.001)	0.003	(0.007)
σ_v^2	0.013		0.003		0.003	
σ_u^2	0.025		0.005		0.004	
σ_α^2	–		0.021		0.001	
λ_u	0.647		0.180		0.539	
λ_α	–		0.721		0.095	
DIC	−463.470		−1100.700		−1112.800	

scores. The relevant estimates are σ_v^2, σ_u^2, and σ_α^2, which stand for the variance of random noise, inefficiency, and separable heterogeneity, respectively; their relative importance (expressed as a share of total error variance) is given by λ_u and λ_α. For the SFM, total cost variability is 0.038 ($\sigma_v^2 = 0.013$, $\sigma_u^2 = 0.025$), with most variation in the inefficiency term. According to $\lambda_u = 0.647$, about 65 percent of cost variability is due to inefficiency, a share comparable to the literature cited in the Introduction section. Next, the RIFM with $\sigma_v^2 = 0.003$, $\sigma_u^2 = 0.005$, and $\sigma_\alpha^2 = 0.021$ confirms the existence of separable heterogeneity. It also indicates a reduction of cost variability by 23 percent (from 0.038 to 0.029), due to its ability to capture heterogeneity in the random intercept. Variability due to inefficiency even declines by 80 percent (from 0.025 to 0.005) and from 65 percent to 18 percent in relative terms. Most of the cost variability between Swiss hospitals can now be attributed to unobserved separable heterogeneity ($\sigma_\alpha^2 = 0.021$), accounting for 72 percent of total cost variability. Still, non-separable heterogeneity is likely to exist, biasing coefficients and inefficiency scores.

The RPFM confirms this concern. As to bias, the capital indicator $BEDS$ now has the expected negative sign (although insignificant) while the coefficient of PL attains the high value suggested by the argument that hospital costs are mainly driven by labor cost. As to cost variability, it again decreases markedly by 72 percent (from 0.029 RIFM to

0.008 RPFM). Most of the cost variability that cannot be attributed to casemix – equal to 0.042 and estimated by the residual sum of squares of a single regression with VC as a dependent variable and $PCASES$ as the independent variable – can now be explained by the estimated cost function. Interestingly, while $\sigma_v^2 = 0.003$ and $\sigma_u^2 = 0.004$ change little from RIFM, the variance of separable heterogeneity diminishes drastically from 0.021 to 0.001, with its relative importance falling from 72 percent to a more plausible 9 percent. Consequently, the relevance of inefficiency increases to 54 percent, comparable to the SFM value.

Indeed, the RPFM attributes most of the cost variability to the technology parameters, as can be seen from the covariance matrix of the Wishart distribution in Table 3.3. The variances on the diagonal show that heterogeneity is strongly related to the output indicator $PCASES$ and the structural variable $INSUR$, in spite of the fact that $PCASES$ already adjusts for heterogeneity through a casemix index. At least in the Swiss case, this raises doubts about the relevance of the casemix index used to adjust for cost variability in prospective payment. As to the off-diagonal entries, the negative correlation between $BEDS$ and PL, although insignificant, points to capital and labor being complements in the hospital sector.

Table 3.3: Variance-Covariance Matrix of the Wishart Distribution

	PCASES	*OUTP*	*PL*	*BEDS*	*INTERN*	*SPEC*	*SUB*	*INSUR*
PCASES	0.505*							
OUTP	−0.338	0.372*						
PL	0.257	−0.196	0.297*					
BEDS	−0.216	0.171	−0.132	0.257*				
INTERN	−0.125	0.088	−0.17	0.091	0.362*			
SPEC	−0.074	0.057	−0.035	0.043	0.001	0.132*		
SUB	0.208	−0.144	0.09	−0.091	0.060	−0.018	0.279*	
INSUR	0.389	−0.318	0.267	−0.238	−0.159	−0.081	0.159	0.584*

∗ Significant at the 95% confidence level.

3.3.3 Efficiency Scores

In many countries, public authorities finance major parts of hospital investment and decide about the opening, closing down, and restructuring of hospitals. For these decisions, efficiency scores may provide guidance. However, as shown in Section 3.3.2, accounting for heterogeneity has an impact on the cost variability attributed to inefficiency (σ_u^2, λ_u). Therefore, a comparison of mean efficiency scores, their distribution, and development over time between the three models is of considerable interest.

Efficient hospitals are on the estimated cost frontier ($\widehat{u_{it}} = 0$, $\widehat{eff} = 1$), those with inefficiency above the frontier ($\widehat{u_{it}} > 0$, $\widehat{eff} < 1$). Since the u_{it} are in logarithms, one has

$$\widehat{eff_{it}} = exp(-\widehat{u_{it}}), \tag{3.14}$$

with $\widehat{u_{it}}$ simulated from a half normal distribution.

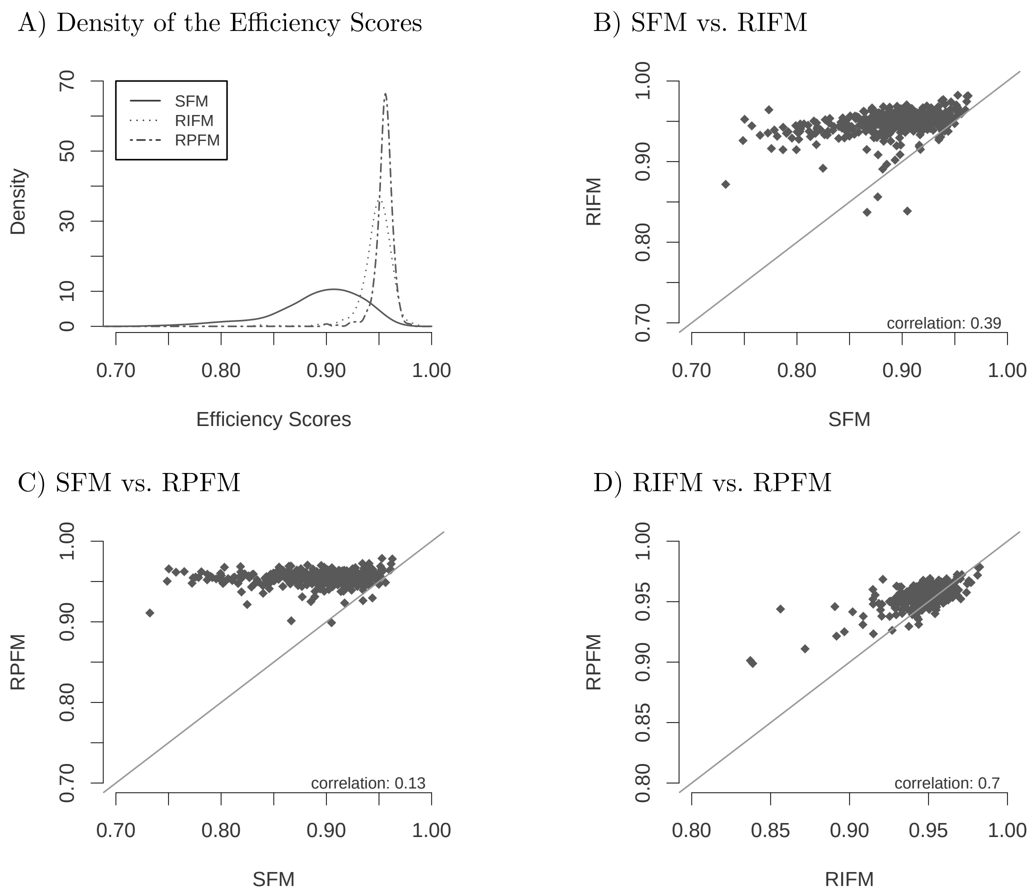
Table 3.4: Efficiency Values by Model Type and Year

	Average	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis
SFM	0.89	0.96	0.73	0.04	−1.07	4.23
RIFM	0.95	0.98	0.84	0.02	−2.78	17.70
RPFM	0.95	0.98	0.90	0.01	−2.06	12.84
2007						
SFM	0.87	0.94	0.75	0.04	−0.90	3.37
RIFM	0.95	0.97	0.84	0.02	−3.74	20.56
RPFM	0.95	0.97	0.90	0.01	−3.09	16.53
2006						
SFM	0.88	0.95	0.75	0.04	−0.90	3.69
RIFM	0.95	0.97	0.86	0.02	−2.69	15.55
RPFM	0.96	0.97	0.94	0.01	−0.54	3.57
2005						
SFM	0.90	0.96	0.73	0.04	−1.46	6.25
RIFM	0.95	0.98	0.87	0.01	−1.77	11.61
RPFM	0.95	0.98	0.91	0.01	−1.35	8.60
2004						
SFM	0.91	0.96	0.78	0.03	−1.41	5.76
RIFM	0.95	0.98	0.89	0.01	−0.96	5.77
RPFM	0.95	0.97	0.92	0.01	−1.13	5.42

Starting with the top of Table 3.4, the first thing to note is that mean efficiency scores are 0.89 or higher, putting the potential for cost reduction at 11 percent or less. This figure is much closer to the 14 percent reported by Hollingsworth (2008) for other European countries and cited in the Introduction section. However, efficiency scores derived from the SFM are markedly lower than their RIFM and RPFM counterparts. Unobserved heterogeneity therefore does lower estimated efficiency scores, causing the potential for cost reduction to be overstated. Over the observation period, the average SFM score is

0.89, suggesting a cost reduction potential of 11 percent. Using the RIFM that corrects for separable heterogeneity, one arrives at a mean score of 0.95, a value comparable to Farsi et al. (2008) who estimated a true random effects model. Thus, prospective payment based on a SFM would overestimate the potential for cost reduction by no less than 6 percent, causing financial distress to at least some cost-efficient hospitals who happen to be stuck with inferior technology e.g. due to old buildings. Turning to the RPFM, which distinguishes non-separable from separable heterogeneity, one does not find a change away from RIFM mean scores. However, the minimum value is now 0.90 rather than 0.84, accompanied by a decrease in (negative) skewness and kurtosis.

Figure 3.1: Efficiency Estimates of the SFM, RIFM, and RPFM, Years 2004-7



While these differences are evident from panel A of Figure 3.1, a comparison of individual efficiency scores is even more telling. Panel B of Figure 3.1 reveals that hospitals that would have been rated below 85 percent efficiency according to SFM gain up to

12 percentage points when the RIFM is applied instead. Panel C shows that this gain may even reach 15 points when the more general RPFM is used. Finally, the comparison between RIFM and RPFM in panel D of Figure 3.1 indicates that hospitals with a RIFM score below 0.92 still would benefit from a transition to RPFM, although the gain rarely exceeds 5 percentage points. Therefore, at a given point of time and for a majority of Swiss hospitals, it clearly matters whether or not unobservable heterogeneity is taken into account in performance measurement.

Still, the three models might agree when it comes to development over time. Returning to Table 3.4, one notices that mean SFM efficiency scores have decreased over time, from 0.91 in 2004 to 0.87 in 2007. In sharp contrast, the RIFM and RPFM measures remained constant. Under the impression of SFM estimates, regulators would therefore have concluded that prospective payments should be cut to squeeze increasingly important cost reductions out of the hospital sector. Yet the evidence points to an increased importance of unobservable heterogeneity (see also the hikes in skewness and kurtosis of RIFM and RPFM scores especially in 2007). Such an increase is credible in view of the fact that in response to sluggish economic growth, hospital renovation projects were postponed or downsized. Failure to control for unobservable heterogeneity thus risks to punish more and more harshly those hospitals that are hampered by outdated technology.

3.4 Concluding Remarks

With prospective reimbursement of hospitals becoming increasingly common, measurement of hospital performance has been gaining in importance. However, measurement can be biased if it fails to account for technological heterogeneity, part of which is unobservable. This paper seeks to shed some light on the importance of both observable and unobservable heterogeneity by estimating a standard cost frontier model, a random intercept cost frontier model, and a random parameter cost frontier model. Bayesian inference is applied to the data of about 100 Swiss hospitals between 2004 and 2007.

Results confirm the existence of unobserved heterogeneity causing some of the cost variability even though inpatient cases are casemix-adjusted. This means that the standard frontier model, which does not control for heterogeneity, is insufficient for deriving unbiased performance measures. In the case of Switzerland, the biases may be substantial. Whereas Hollingsworth (2008) reports a potential for cost reductions of 14 percent for Europe, Farsi et al. (2006) put it to 20 percent for Swiss hospitals covering the years 1997 to 2002. In this paper, a comparable standard frontier model suggests 11 percent for 2004 to 2007, dropping to 5 percent when both observable and unobservable heterogeneity are controlled for. An element-wise comparison reveals that hospitals rated 85 percent efficient and less (using the standard method) would gain up to 12 percentage points. Therefore, quite a few hospitals, although highly efficient in fact, would end up in financial distress if regulators were to cut reimbursement rates in an attempt to enforce the cost reductions indicated by the standard frontier model. This underlines the importance of accounting for both observable and unobservable heterogeneity in the estimation of hospital cost frontiers.

However, this analysis is not without limitations. First, there is the risk of misspecification causing bias in the benchmark values used to characterize the posterior distribution governing Bayesian inference. Specifically, the cost of capital is missing from the equation; but in addition, the indicator for unit cost of labor is an aggregate over skill categories ranging from physicians to orderlies, and just counting the categories of internships offered likely constitutes a poor measure of educational services provided. Second, the distinction between observable and unobservable heterogeneity remains somewhat arbitrary; for instance, if measures of the vintage of hospital capital stock in terms of buildings and medical technology were available, a greater part of cost variability would be attributed to the observable component, likely causing the estimated influence of total heterogeneity on cost to be reduced. This point relates to a third weakness, which is that management inefficiency continues to be measured as a residual rather than by direct indicators. Therefore, by minimizing the contribution of this residual to cost variation, random parameter models might end up going too far in exonerating hospital management.

This said, the evidence presented here does suggest that the standard cost frontier model is insufficient for measuring hospital performance due its failure to take technological heterogeneity into account at all. While the evidence is limited to a sample of Swiss hospitals, the reasons for heterogeneity are of a general nature. They also apply to other heavily regulated or public sectors such as energy, education, and banking, underlining the importance of specifying cost frontier models that yield unbiased efficiency scores.

Appendix

Table 3.5: Econometric Results of the Random Technology Parameters

	RIFM	RPFM								
	CONSTANT	*CONSTANT*	*PCASES*	*OUTP*	*PL*	*BEDS*	*INTERN*	*SPEC*	*SUB*	*INSUR*
$\bar{\bar{\beta}}$	3.106	4.948	0.485	0.013	0.384	−0.148	0.008	0.000	0.231	0.003
	(0.32)	(1.11)	(0.115)	(0.011)	(0.110)	(0.127)	(0.006)	(0.007)	(0.529)	(0.007)
$K111$	0.693	0.439	0.170	−0.157	−0.013	−0.023	0.007	0.014	0.480	0.000
	(0.14)	(3.04)	(0.780)	(0.317)	(0.745)	(0.935)	(0.019)	(0.017)	(3.008)	(0.051)
$K112$	0.273	0.865	−0.250	0.066	−0.166	0.282	−0.004	0.002	0.915	−0.008
	(0.07)	(2.56)	(0.190)	(0.096)	(0.300)	(0.41)1	(0.010)	(0.013)	(2.552)	(0.014)
$K121$	0.187	0.992	−0.231	0.028	−0.295	0.241	−0.007	0.010	1.024	0.007
	(0.06)	(2.57)	(0.251)	(0.096)	(0.254)	(0.394)	(0.006)	(0.011)	(2.574)	(0.013)
$K122$	0.097	0.105	−0.220	−0.012	−0.183	0.474	−0.005	0.000	0.601	0.002
	(0.04)	(1.58)	(0.157)	(0.012)	(0.145)	(0.205)	(0.008)	(0.009)	(0.955)	(0.010)
$K231$	0.082	0.676	−0.247	0.151	−0.129	0.046	−0.010	−0.005	0.682	0.002
	(0.07)	(2.16)	(0.236)	(0.105)	(0.276)	(0.229)	(0.080)	(0.019)	(0.987)	(0.014)
M_{05}	0.034	0.462	0.078	−0.004	−0.148	−0.065	0.000	0.000	−0.054	0.000
	(0.01)	(0.33)	(0.047)	(0.005)	(0.077)	(0.047)	(0.001)	(0.001)	(0.044)	(0.001)
M_{06}	0.129	1.148	−0.074	0.005	−0.152	0.069	0.000	−0.001	−0.071	0.000
	(0.01)	(0.43)	(0.059)	(0.009)	(0.105)	(0.063)	(0.001)	(0.001)	(0.062)	(0.001)
M_{07}	0.155	0.902	−0.130	0.004	−0.041	0.108	0.000	0.000	−0.057	0.000
	(0.01)	(0.45)	(0.062)	(0.010)	(0.108)	(0.064)	(0.001)	(0.001)	(0.064)	(0.001)

CHAPTER IV

Does Prospective Payment Increase Hospital (In)Efficiency? Evidence from the Swiss Hospital Sector

Philippe K. Widmer

Abstract: Several European countries have followed the United States in introducing prospective payment for hospitals with the expectation of achieving cost efficiency gains. This article examines whether theoretical expectations of cost efficiency gains can be empirically confirmed. In contrast to previous studies, the analysis of Switzerland provides a comparison of a retrospective per diem payment system with a prospective global budget and a payment per patient case system. Using a sample of approximately 90 public financed Swiss hospitals during the years 2004 to 2009 and Bayesian inference of a standard and a random parameter frontier model, cost efficiency gains are found, particularly with a payment per patient case system. Payment systems designed to put hospitals at operating risk are more effective than retrospective payment systems. However, hospitals are heterogeneous with respect to their production technologies, making a random parameter frontier model the superior specification for Switzerland.[7]

Keywords: hospital inefficiency, prospective payment system, Bayesian inference, stochastic frontier analysis

JEL classification: C11; C23; D24; I18

---

[7] Author's Note:The author would like to thank Alexander Ziegenbein for his efforts in the data collection.

4

Does Prospective Payment Increase Hospital (In)Efficiency? Evidence from the Swiss Hospital Sector

4.1 Introduction

Growing health care expenditures over the last decades have highlighted the need for health care reforms in order to contain future cost increases. One promising approach, which was first implemented in the U.S. and was recently adapted by many European countries, involves the transition from retrospective (RPS) to prospective (PPS) hospital payment systems (see Smith, 2004 and Schneider, 2007 for an overview of Europe's reforms). The assumption is that a change to predetermined and fixed payments would place hospitals at operating risk and would increase their cost efficiency.

Even though there are convincing theoretical arguments for cost reductions and efficiency gains (Biorn et al., 2003, Chalkley and Malcomson, 1998, and Newhouse, 1996), empirical literature is lacking. The linkage between efficiency gains and PPS has yet to be demonstrated for the U.S. Medicare reform of 1983, which switched from RPS to a payment per patient case system, or for any of the European countries that moved from RPS to a payment per patient case or a global budget system (see Section 4.2 for further information on the payment systems). For example, Borden (1988) found no significant efficiency gains for 93 New Jersey hospitals from the years 1979 to 1984. Similar results were obtained by Chern and Wan (2000) when they examined the catch-up effect of technically inefficient hospitals in Virginia from 1984 to 1993. Inefficient hospitals became even more inefficient in 1993, which is contrary to the expectations of PPS. However, efficiency gains were shown by Morey and Dittman (1996), who analyzed the technical inefficiency of 105 hospitals in North Carolina. The results of European reforms remain inconclusive. While no efficiency gains were found in Austrian hospitals after funding shifted from per diem payments to global budgets in 1997 (Sommersguter-Reichmann, 2000), gains were found in Portugal (Dismuke and Sena, 1999), Finland (Linna, 1999), and Norway (Biorn et al., 2006). Thus, hospital costs could even increase with PPS. Since PPS is well known to be concurrent to higher administration and supervising costs, which is not yet included in theoretical models, the incentive for cost reduction could be overstated.

However, these inconclusive results are most likely due to a lack of analytical rigor. In particular, although it is widely accepted that hospitals are rather heterogeneous in their production of health care services (Widmer et al., 2011), previously applied Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) have been restricted to homogeneous technology. Furthermore, it is well known that results of the frequently applied two-stage DEA approach are biased since it does not account for a possible correlation of the independent variables with the inputs and outputs of the first-stage DEA (Simar and Wilson, 2000). Finally, since most countries only recently switched to PPS at the country level, the time series available for within treatment analysis have been very short. Most studies analyze a time period of four to five years, which may be too short for any reliable conclusion to be drawn. For instance, any changes could be driven by unobserved exogenous shocks, such as new medical technologies or inflation, that occurred concurrent with the implementation of PPS (Linna, 1999).

In order to overcome the limitations of previous studies, this article compares a retrospective per diem system with a prospective global budget and a payment per patient case system using data from Switzerland, where some member states (cantons) changed to different variants of PPS while others remained with RPS. The contribution of this article is twofold. First, it extends previous work by implementing a random parameter frontier model to control for the importance of unobserved heterogeneity among six hospital categories and addresses whether empirical results significantly depend on the assumptions made for the production technology. Second, it determines whether theoretical expectations for cost savings can be confirmed in empirical analysis by relating calculated inefficiency scores to the three payment systems. Estimates are derived by an extended single-step approach of Battese and Coelli (1995).

The empirical analysis reveals two key results. First, with respect to model comparison, the random parameter frontier model is more robust and has a higher explanatory power than the single cost frontier model. Heterogeneity correction among hospital categories is crucial in deriving meaningful inefficiency scores. Second, PPS are negatively correlated with hospital cost inefficiency, particularly the payment per patient case system. Payment systems designed to put hospitals at operating risk are more effective than retrospective payment systems in containing hospital costs.

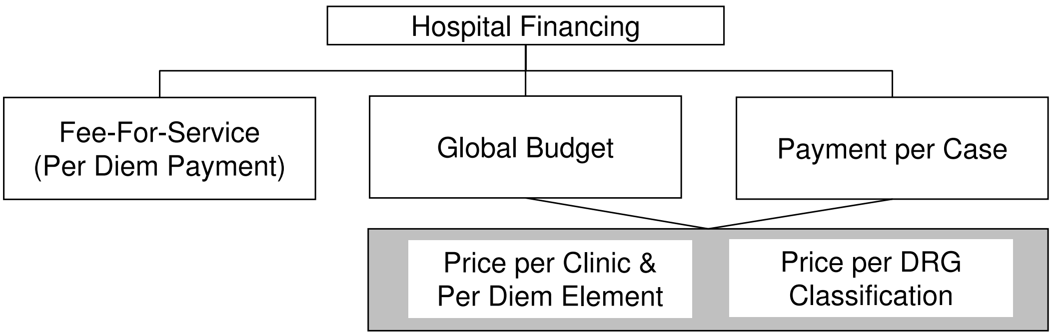
The remainder of this paper is structured as follows. Section 4.2 gives an overview of the different prospective and retrospective payment systems that coexisted in Switzerland between 2004 and 2009. Section 4.3 outlines the standard and random parameter frontier model and Section 4.4 describes the data used as well as the empirical specifications. Finally, Section 4.5 presents the results of the cost frontier models and the determinants of inefficiency.

4.2 Introduction to Swiss Hospital Financing

The Swiss health care system has been shaped by the country's decentralized federal structure, in which all 26 cantons are responsible for providing health care services to their residents. The hospital sector is no exception. Cantonal authorities are responsible for capacity planning and for the quality of hospital care. Provision is typically purchased from hospitals that are qualified to provide health care to primary insured patients. However, this does not imply that hospital financing only comes from cantonal sources. On the contrary, health insurers pay an agreed amount of up to 50 percent of operating costs, resulting in a dual system where cantons cover the residual cost and investments in infrastructure.[1] Modes of financing can therefore differ among cantons.

Increasing health care costs have induced many cantons to revise their hospital payment system. Especially the implementation of the new federal law on social health care insurance in 1994 (effective in 1996), where cantonal authorities were given legislative power to control for hospital operating costs, has resulted in the coexistence of various RPS and PPS in Switzerland (see Figure 4.1 for an overview of existing payment systems).

Figure 4.1: Swiss Payment Systems



Prior to 1996, cantons primarily used a retrospective cost-based per diem or fee-for-service system to pay hospitals for their services. Remuneration was equal to reported costs and bankruptcy was only possible if cantonal authorities decided to reduce overcapacity. Unsurprisingly, critics of these schemes argued that there was little incentive for cost containment. Hospitals could waste resources and increase health care costs in order to obtain greater reimbursement. Hence, after 1996, several cantons experimented with PPS to set incentives for cost containment. The two alternatives included a global budget and a payment per patient case system (see second level in Figure 4.1). Under a global budget system, hospitals are paid a fixed amount for a predetermined number of admissions whether or not a patient seeks care during the accounting period. Under a payment per patient case system, hospitals are paid a fixed amount per admission, regardless of the actual cost. In both cases, hospitals obtain the gain or incur the loss,

[1] Health care insurers might cover more than 50 percent of expenditures in privately owned hospitals, which are not on the cantonal list. However, these are typically for-profit hospitals specializing in supplementary insured patients.

making them act to minimize costs. However, the incentive for cost minimization could be weakened in the Swiss case because many cantons still do not firmly exclude a bailout. This is especially the case in a global budget system, where hospitals are generally allowed to renegotiate their budget for unexpected costs. Hospitals still have an implicit deficit guaranty which reduces their operating risk and therefore the incentive for cost minimization. Furthermore, the determination of the remuneration per admission could also influence incentives for cost containment. Two variants are widely used in Switzerland (see third level in Figure 4.1). The first variant determines payments per admission according to a clinic-specific average price and a per diem element to control for differences in the length of stay. The second variant uses a Swiss specification of the DRG[2] classification system that attempts to classify patients into groups with similar usage of resource. In contrast to the first variant, payments are independent of the length of stay. Thus, hospitals have no incentives to maximize the length of stay, which should result in additional cost savings. Even in a DRG system there is provision for additional payment for those patients who are unusually expensive within the DRG classification, but these outlier payments apply to only a small portion of patients and are not directly related to length of stay.

An increasing number of cantons have changed to PPS. In 2004, only 38 percent of all Swiss hospitals were still reimbursed by per diem payments. Most hospitals had PPS and almost 36 percent of them already used DRG classifications. In 2007, the number of hospitals with PPS increased even more and most cantons used DRG classifications (see Meister, 2008). Unsurprisingly, in 2007 the Swiss parliament revised the insurance law to introduce a DRG system in all cantons by 2012. Following the U.S. Medicare reform of 1983 and the German reform of 2004, policy makers believe that the new reimbursement system would increase cost efficiency as outlined in the Introduction section. This article aims to determine whether the DRG system is preferable to contain health care costs. The hypotheses of interest are:

[2] Diagnostic Related Group (DRG) is a diagnoses classification system, to distinguish between different requirements on hospital sources. DRG has first been used in the US health care program for elderly people (Medicare,1983). In Switzerland, AP-DRG – a non-profit organization – so far provided information to different diagnostic groups and their cost weights. In 20012, SwissDRG is going to replace the AP-DRG system.

1. Hospitals with PPS are more cost efficient than hospitals with RPS. Putting a hospital at any amount of operating risk should strengthen incentives for cost minimization (lower cost inefficiency).

2. Hospitals with payments based on DRG classifications are more cost efficient than those paid with a per diem element. The fact that DRG systems do not account for longer length of stay should cause additional cost savings (lower cost inefficiency).

4.3 Estimation Models

In order to analyze these hypotheses, firm-specific inefficiency scores must first be established from estimated cost frontiers.[3] This paper applies two specifications to check for the importance of unobserved heterogeneity among hospitals. The first specification is a standard frontier model that was first implemented by Aigner et al. (1977) and Meeusen and van den Broeck (1977). It estimates inefficiency as the distance between a cost frontier and observed expenditures. Observable heterogeneity is captured by shifting means of the inefficiency term, similar to preliminary work by Battese and Coelli (1995), Huang and Liu (1994), and Kumbhakar et al. (1991). The second specification is a random parameter frontier model that additionally controls for unobserved heterogeneity in technology parameters (for other applications see Orea and Kumbhakar, 2004 and Tsionas, 2002). Inefficiency is estimated as cost deviations from category-specific cost frontiers.

4.3.1 The Standard Frontier Model (SFM)

The cost frontier for hospital $i = 1, ..., N$ at time period $t = 1, ..., T$ can be written as

$$C_{it} = C(Y_{it}, W_{it}; \alpha, \beta) + \overbrace{u_{it} + v_{it}}^{\varepsilon_{it}}, \qquad (4.1)$$

with C_{it} representing operating expenditures, Y_{it} denoting the output vector, and W_{it} as the vector of input prices. α is the intercept and β is a $(K \times 1)$ vector of unknown slope parameters. $C(Y_{it}, W_{it}; \alpha, \beta)$ is the deterministic part of the cost frontier that remains to

[3] See Coelli et al. (2005) and Kumbhakar and Lovell (2000) for an overview of inefficiency measurement methods.

be specified for the empirical estimation. Typically, this is either a Cobb-Douglas or a more flexible translog functional form.

The error term ε_{it} is split into two additive components, enabling deviations for random noise, v_{it} and cost inefficiency, u_{it}. Random noise is normally distributed $v_{it} \overset{iid}{\sim} N[0, \sigma_v^2]$ with mean zero and variance σ_v^2. Firm-specific inefficiency u_{it} is assumed to follow a one-sided distribution supported on the interval $[0, \infty)$. The larger u_{it}, the more cost inefficient a hospital and the greater the potential for cost savings.

Since the main purpose of this paper is to analyze the influence of PPS on inefficiency, inefficiency is specified congruent to Battese and Coelli (1995) as a truncated normal distribution $u_{it} \sim f_{N^+}[\bar{u}_{it}, \sigma_u^2]$ with firm specific means $\bar{u}_{it}$ and variance σ_u^2. In this article, mean inefficiency is a linear function of $l = 1, .., L$ explanatory variables Z_{it} that influence inefficiency,

$$\bar{u}_{it} = \gamma_o + \sum_{l=1}^{L} \gamma_l Z_{lit} + \varsigma_{it}, \tag{4.2}$$

where γ is an $(L \times 1)$ vector of unknown parameters to be estimated and ς_{it} remains as unexplained hospital-specific inefficiency.

4.3.2 The Random Parameter Frontier Model (RPFM)

One way to extend the SFM for unobserved heterogeneity is the random parameter frontier model, which estimates inefficiency scores from individual cost frontiers.[4] In this article, the model accounts for $j = 1, ..., J$ exogenously given hospital categories that are expected to have different production technologies,

$$\begin{aligned} C_{it} &= C(Y_{it}, W_{it}; \alpha_j, \beta_j) + \overbrace{u_{it} + v_{it}}^{\varepsilon_{it}}, \\ u_{it} &\sim f_{N^+}\ [\bar{u}_{it}, \sigma_u^2], \qquad \bar{u}_{it} = \gamma_o + \sum_{l=1}^{L} \gamma_l Z_{lit} + \varsigma_{it}, \\ \alpha_j &= \alpha + w_j, \\ \beta_j &= \beta + w_j. \end{aligned} \tag{4.3}$$

[4] It is worth mentioning that a separation is not preferable in every case. If technology is manageable than the choice of an inferior technology could be treated as inefficiency.

Different from SFM, this specification allows inefficiency to be disentangled from unobservable heterogeneity with category-specific intercepts $\alpha_j = \alpha + w_j$ and slope parameters $\beta_j = \beta + w_j$. All time-invariant and firm-specific heterogeneity is captured in w_j, which is a $[(K+1) \times 1]$ vector of random variables.

This paper specifies α_j and β_j similar to Tsionas (2002) as a multivariate normal distribution

$$\begin{pmatrix} \alpha_j \\ \beta_j \end{pmatrix} \sim f_{MN} \left[\begin{pmatrix} \bar{\alpha} \\ \bar{\beta} \end{pmatrix}, \Sigma \right], \quad \text{with } \Sigma \sim f_W \begin{bmatrix} \sigma_\alpha^2 & \sigma_{\alpha,\beta} \\ \sigma_{\alpha,\beta} & \sigma_\beta^2 \end{bmatrix}, \tag{4.4}$$

where $\bar{\alpha} \sim N[0, \sigma^2_{\bar{\alpha}}]$ and $\bar{\beta} \sim N[0, \sigma^2_{\bar{\beta}}]$ are both normally distributed with mean zero and variance (σ_u^2, σ_v^2). This is a hierarchical model that first measures the mean effects $(\bar{\alpha}, \bar{\beta})$ and then estimates individual effects (α_j, β_j) for each parameter. Σ is Wishart distributed with a $[(K+1) \times (K+1)]$ positive definite covariance matrix $S = (\sigma_\alpha^2, \sigma_\beta^2, \sigma_{\alpha,\beta})$, denoting unobserved heterogeneity among hospitals. For $\Sigma = 0$ no variation exists and the RPFM simplifies to a SFM.

Based on the distributional assumptions made in the SFM and RPFM, Bayesian econometrics is applied for the simultaneous estimation of the parameters in the cost frontier and the inefficiency term. This is superior to the frequently applied classical maximum likelihood statistics since it considers unknown parameters as random variables, specified as prior distributions. Exact small sample results are possible because of the prior information included. Estimation is performed using R and WINBUGS. Corresponding Bayesian specifications and programming codes are described in the Appendix.

4.4 Data and Econometric Specifications

4.4.1 The Sample

Data used in this study were provided by the annual reports of the federal office of public health and by the conference of cantonal health ministers. They include 333 Swiss hospitals for the time period of 2004 to 2009, consisting of 5 university hospitals, 23 central hospitals, 27 large regional hospitals, 46 medium regional hospitals, 46 small regional hospitals, 28 specialized surgery hospitals, and sundry hospitals, viz. psychiatric and rehabilitation clinics. In total, 127 of the 333 hospitals are private and not subsidized.

In the interest of comparability, the entire data set was reviewed and assessed for the presence of any missing data and outliers that could distort the results. Furthermore, the sundry category and all non-subsidized hospitals were discarded. An unbalanced panel consisting of 545 observations from six different hospital categories with sufficient quality was finally analyzed. In Table 4.1, the variables are listed together with descriptive statistics.

Table 4.1: Definition and Descriptive Statistics of the Variables Used in the Analysis

Variable	Definition	Mean	Min	Max
VC	Variable operational expense (VC)[1)]	135,621	8,550	1,015,756
Y_1	No. of inpatient cases, CMI-adjusted ($CASES$)	9,113	502	52,143
Y_2	Revenue from outpatients ($OUTP$)[1)]	27,418	0	223,937
PL	Labor input price (PL)[1)]	101	34	146
PM	Price of other production inputs (PM)[1)]	4	2	7
K	No. of beds ($BEDS$)	229	31	1,169
S_1	No. of internship categories ($INTERN$)	22	0	134
S_2	No. of specialties ($SPEC$)	39	4	106
Z_1	Dummy= 1 for prospective payment systems (PPS)[2)]	78	0	100
Z_2	Dummy= 1 for payments per patient case ($CASEP$)[2)]	14	0	100
Z_3	Dummy= 1 for global budgets ($GLOB$)[2)]	64	0	100
Z_4	Dummy= 1 for DRG classifications (DRG)[2)]	51	0	100
T_t	Year dummies, $t = 2005$ to 2009 (base year is 2004)			

[1)] in 1,000 CHF, 1 CHF=0.8 USD (2004 exchange rates).
[2)] in percent, PPS=78 in column three means that on average 78 percent of all hospitals have PPS.

4.4.2 Specification of the Cost Frontier

With these data, a variable Cobb-Douglas cost frontier (subscripts $i = 1, ..., N$ and $t = 1, ..., T$ are dropped for simplicity) can be specified as

$$ln\frac{VC}{PM} = \alpha + \sum_{m=1}^{2} \beta_m lnY_m + \beta_3 ln\frac{PL}{PM} + \beta_4 lnK + \sum_{l=1}^{2} \beta_l S_l + \sum_{t=1}^{5} \beta_t T_t + u + v, \quad (4.5)$$

where variable cost (VC) depends on two output categories (Y), one input price for labor (PL), one price for other production inputs (PM), one capital stock (K), two structural variables (S), and five time dummies (T) to control for any unobserved dynamics over time (base year 2004). Normalizing VC and PL by PM imposes linear homogeneity in input prices.

Health care output – change in health status – is difficult to measure directly for Swiss hospitals. In this article, measures for inpatient care $CASES$ and outpatient care $OUTP$ serve as intermediate outputs. To adjust for severity in inpatient care, CMI-adjusted admissions are used. Outpatient care is approximated by ambulatory earnings, similar to Farsi et al. (2006) and Biorn et al. (2003). Furthermore, input price PL is calculated as labor expense divided by the number of full time employees. Input price PM aggregates all the remaining inputs, such as energy, material, and purchased services that cannot be distinguished due to data limitations. An approximate price for PM is calculated as residual cost divided by the number of admissions (a discussion of this common simplification is given in Coelli et al., 2005, ch. 5). Since capital stock (total fixed assets) is hardly measurable, $BEDS$ serve as an approximation. Finally, the number of internship categories $INTERN$ and specialties $SPEC$ control for observable service heterogeneity among hospitals.

The formulation can be justified on several grounds. First, it is compatible with short-term cost minimization, reflecting the fact that capital (indicated by $BEDS$) is a pre-determined rather than a decision variable. In Switzerland, cantonal hospital planning divisions mainly decide capacity. Second, the exclusion of user cost of capital from the equation avoids measurement errors since values would have to be imputed because most hospitals are not charged capital user costs.

4.4.3 Determinants of Inefficiency

Since the influence of PPS on inefficiency is the focus of this article, additional explanatory variables are included in the inefficiency term – see eq. (4.2) – to test for the two hypotheses from Section 4.2:

(1) Hospitals with PPS are more cost efficient than hospitals with RPS;

(2) Hospitals with payments based on DRG classifications are more cost efficient than those paid with a per diem element.

Hypothesis (1) is tested with two models. Model (1) refers to eq. (4.6),

$$\bar{u}_{it} \quad = \quad \gamma_o + \gamma_1 PPS + \gamma_2 PPS{:}DRG + \varsigma_{it}, \tag{4.6}$$

which relates mean inefficiency to a dummy variable that equals one for hospitals with PPS and zero for hospitals with RPS. It determines whether PPS – either a global budget or a payment per patient case system – is more effective than the retrospective alternative.

Model (2) refers to eq. (4.7),

$$\bar{u}_{it} \quad = \quad \gamma_o + \gamma_1 CASEP + \gamma_2 GLOB + \gamma_3 CASEP{:}DRG + \gamma_4 GLOB{:}DRG + \varsigma_{it}, \tag{4.7}$$

which is a refinement of Model (1). It checks for the unique effects of a global budget and a payment per patient case system. Therefore, the variable PPS is replaced by two dummy variables, $GLOB$, for hospitals with a global budget, and $CASEP$, for hospitals with payments per patient case. Hospitals with a retrospective per diem system form the control group. In Model (2), it is expected (from Section 4.2) that hospitals receiving payments per patient case are more cost efficient (have lower inefficiency scores) since most hospitals with global budgets have a partial deficit guaranty.

Hypothesis (2) calls for the introduction of an additional dummy variable, DRG, in eqs. (4.6) and (4.7), which is specified as a nested interaction term. It measures the supplementary effect of DRG classifications relative to the alternative specification with a per diem element. As outlined in Section 4.2, payments based on a per diem element can reduce incentives for cost minimization since hospitals have incentive to increase the length of stay. Therefore, it is expected that hospitals with DRG, which is free from any adjustment for length of stay, are more efficient (have lower inefficiency scores) than the frequently applied alternative.

Additionally, Model (3)checks for the hypothesis (2) with eq. (4.8),

$$\bar{u}_{it} \quad = \quad \gamma_o + \gamma_1 PPS + \gamma_2 PPS{:}DRG + \gamma_3 PPS{:}DRG_1 + \gamma_4 PPS{:}DRG_2 + \varsigma_{it}. \tag{4.8}$$

This is a refinement of Model (1) for a possible catch-up effect of DRG over time. Therefore, two additional dummies DRG_j, $j = 1, 2$ are included, where j indicates the time lag from the initiation of the reimbursement scheme. Since it is possible that DRG only becomes effective a few years after initiation, it is preferable to test for these effects as well.

4.5 Empirical Results

This section first presents estimates of the technology parameters and inefficiency scores with a special focus on the influence of unobserved heterogeneity. Second, in Section 4.5.2, the influence of PPS is discussed for the three models outlined in Section 4.4.3.

To obtain posterior estimates, Monte Carlo Markov Chain (MCMC) algorithms were run for 20,000 iterations and the first 10,000 samples were discarded as a burn-in phase. Different assumptions for priors and starting values converged to roughly the same values without strong periodicities or tendencies in the trace plot. Furthermore, the Monte Carlo error is very low. All cost frontier parameters and inefficiency scores have a Monte Carlo error lower than $7.02 * 10^{-4}$, indicating that the results are quite precise and have reached the equilibrium distribution.

4.5.1 Cost Frontier Estimates and Inefficiency Scores

Table 4.2 shows the estimates of the technology parameters of the variable cost frontier from eq. (4.5) after an analysis of cost drivers together with tests for endogeneity, heteroscedasticity, and the skewness of the composite error term were performed. Hospital output could be endogenous in the RPS when hospitals have incentive to increase their output due to higher remuneration. However, a Hausman test did not suggest rejection of the exogeneity assumption. Heteroscedasticity was also not a problem according to a Breusch-Pagan test. Only $INTERN$ had a weak effect on the variance of the composite error term. Finally, because inefficient hospitals by definition lie above the cost frontier, a positively skewed composite error term is required for efficiency measurement. Otherwise, no inefficiency would exist and OLS would be sufficient to estimate the cost frontier. However, residuals of the cost driver analysis were positively skewed, indicating that inefficiency does exist in the Swiss hospital sector.

In the SFM, the first three columns of Table 4.2 contain the estimation mean, the 2.5, and 97.5 percentile of the technology parameters. They satisfy economic conditions in that the cost frontier monotonically increases in the outputs $CASES$ and $OUTP$ as well as in the input price PL. The only exception is hospital beds ($BEDS = 0.22$). Since it is an indicator of capital stock, it should have a negative sign (see Kumbhakar and Lovell, 2000). However, because hospital capacity (no. of hospital beds) is exogenously determined by the cantonal authority, the expected substitution effect could be very small. Furthermore, because $BEDS$ is a poor proxy for capital stock, which is highly correlated with hospital output, estimates might show an output rather than a substitution effect (see e.g. Filippini et al., 2004 for similar difficulties). Consequently, an investigation of the shadow price is not possible, but inefficiency scores are still derivable. Moreover, variable costs tend to shift up systematically over time, with a maximum in 2009 ($T_{09} = 0.036$). Finally, with regard to service heterogeneity it is not surprising that internship categories ($INTERN = 0.002$) and the number of specialties ($SPEC = 0.001$) have a positive effect. Variable cost increases with the number of different services offered.

Table 4.2: Econometric Results for the SFM and RPFM, Years 2004-09

Variables[1]	SFM			RPFM			
	Mean	2.50%	97.50%	Mean	2.50%	97.50%	SD[2]
$Constant$	-0.088	-0.103	-0.073	-0.123	-0.201	-0.059	0.074
$CASES$	0.744	0.708	0.781	0.592	0.359	0.768	0.224
$OUTP$	0.004	-0.002	0.011	0.076	-0.045	0.217	0.158
PL	0.382	0.347	0.418	0.434	0.286	0.612	0.169
$BEDS$	0.220	0.182	0.258	0.281	0.118	0.482	0.159
$INTERN$	0.002	0.001	0.002	0.000	-0.001	0.002	0.001
$SPEC$	0.001	0.000	0.002	0.001	0.000	0.002	0.001
T_{09}	0.036	0.008	0.065	0.017	-0.014	0.045	0.019
T_{08}	0.023	-0.005	0.051	0.007	-0.013	0.031	0.009
T_{07}	0.012	-0.016	0.040	0.002	-0.016	0.024	0.008
T_{06}	0.023	-0.005	0.051	0.009	-0.011	0.032	0.009
T_{05}	0.009	-0.017	0.034	0.004	-0.015	0.028	0.011
σ_u	0.013	0.009	0.017	0.001	0.000	0.003	
σ_v	0.004	0.003	0.006	0.003	0.002	0.004	
DIC	-1106			-1596			
Obs.	545			545			

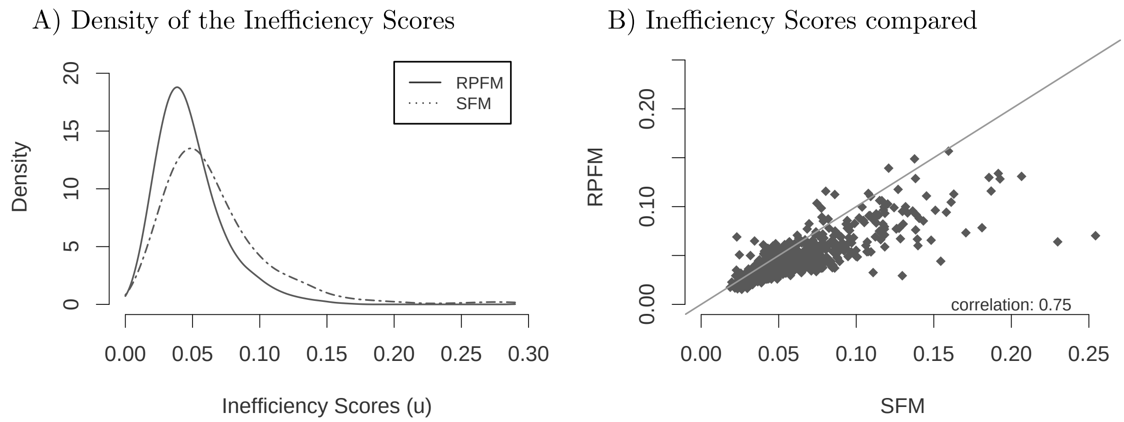
[1] Variable cost (VC) is the dependent variable. Determinants of inefficiency are shown separately in Table 4.3.
[2] SD estimates reveal the diagonal of the covariance matrix Σ of eq. (4.4).

Estimates for the RPFM only have slightly different values [estimation means are represented by $\bar{\beta}$ of eq. (4.4)]. However, the results in the last column suggest that there is a fair amount of variation in the frontier model parameters. Estimates reveal the diagonal of the covariance matrix Σ of eq. (4.4), which can be interpreted as the variation in the parameters across hospital categories. Heterogeneity is highest for inpatient care ($CASES$ = 0.224), followed by the input price for labor (PL = 0.169), capital stock ($BEDS$ = 0.159), and outpatient care ($OUTP$ = 0.158). It is remarkable that even though heterogeneity in inpatient care is already adjusted for by a casemix index, indisputable variation remains among hospital categories (see Section 3.3.1 for similar findings). This raises doubts about the relevance of the DRG classifications to control for cost variability in inpatient care. However, in order to determine whether the greater flexibility of the RPFM is indicated by the data, both models are assessed by the DIC information criteria shown in Table 4.2 (Spiegelhalter et al., 2002). The lower the DIC-value, the better the goodness of fit of the estimated cost frontier, indicating that the RPFM ($DIC = -1596$) has better fit than the SFM ($DIC = -1106$). The SFM seems to be too restrictive for Switzerland. More flexible variants are needed to capture all the existing heterogeneity among hospital categories.

For the present study, the more important question is the impact of unobserved heterogeneity on the estimated inefficiency scores. This is shown in the density and scatter plots in Figures 4.2 and 4.3. Figure 4.2 shows the inefficiency scores of the SFM and the RPFM. Figure 4.3 presents preliminary indications of the influence of PPS.

As shown in Figure 4.2, there are strong differences across the two models. Unobserved heterogeneity increases estimated inefficiency scores, potentially resulting in the overstatement of cost reduction (Figure 4.2 panel A). The mean inefficiency score of the SFM is 0.066, meaning that Swiss hospitals could on average reduce 7 percent of their variable costs. However, using the RPFM, mean inefficiency reduces to about 5 percent. Approximately 2 percent of the SFM scores can be detected as unobserved heterogeneity. A comparison of the individual scores in panel B is even more revealing. Although both models have a high correlation of 0.75, hospitals are systematically measured as more inefficient in the SFM. In particular, hospitals that would have been rated highly inefficient in the SFM gain ground when the RPFM is applied. The maximum inefficiency

Figure 4.2: Estimated Inefficiency Scores of the SFM and RPFM, Years 2004-9



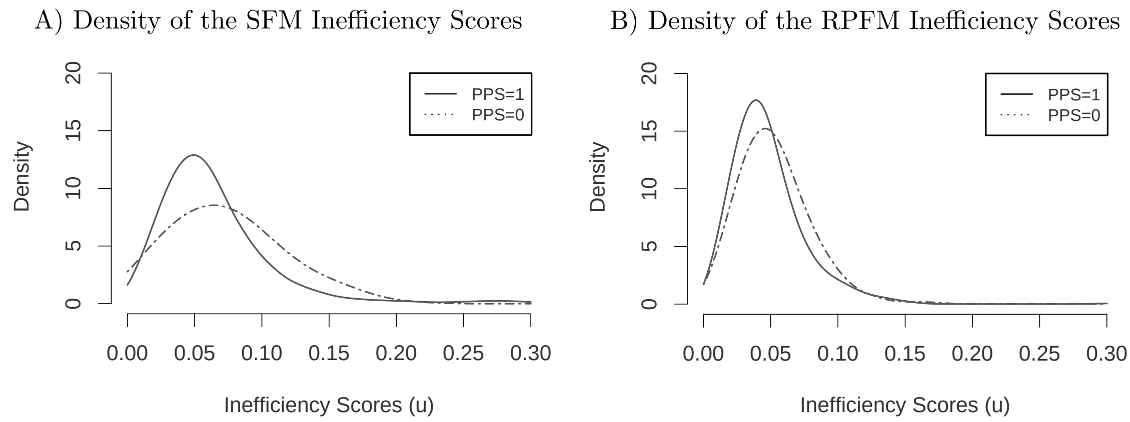
	Mean	Minimum	Maximum	Std. Dev.	Skewness	Kurtosis
SFM	0.066	0.019	0.294	0.039	2.277	10.851
RPFM	0.049	0.015	0.157	0.024	1.417	5.364

score decreases from 0.294 for the SFM to 0.157 for the RPFM, putting the maximum cost reduction at about 16 percent. At a given point in time and for the majority of Swiss hospitals, it clearly matters whether or not unobserved heterogeneity is taken into account.

It is noteworthy that even the SFM reveals significant lower average inefficiency scores than in previous studies to Swiss hospitals (see Widmer et al., 2011, Farsi and Filippini, 2006, and Steinmann and Zweifel, 2003). Compared to Widmer et al. (2011), the SFM inefficiency scores drop from 11 percent to approximately 7 percent. However, since only public financed hospitals are used in this study, differences can be seen as an indication for cost variability between private and public hospitals.

Next, Figure 4.3 reveals some preliminary indications for the effectiveness of PPS. Surprisingly, even though SFM scores are indisputably biased, both models come up with comparable conclusions that PPS reduces hospital cost inefficiency. In the SFM, mean inefficiency decreases from 0.073 to 0.064 (Figure 4.3 panel A). In the RPFM, mean inefficiency scores decrease from 0.053 to 0.047 Figure (4.3 panel B). Both reductions are significant according to a Wilcoxon rank-sum test (the hypothesis that mean inefficiency is equal for the two groups can be rejected at the 95 percent confidence level). However,

Figure 4.3: Estimated Inefficiency Scores by Model Type, Years 2004-9



SFM	Mean	Minimum	Maximum	Std. Dev.
PPS=0	0.073	0.022	0.181	0.036
PPS=1	0.064	0.019	0.294	0.040

RPFM	Mean	Minimum	Maximum	Std. Dev.
PPS=0	0.053	0.017	0.157	0.024
PPS=1	0.047	0.015	0.149	0.023

the decrease in inefficiency is larger in the SFM (mean $= -0.009$) than in the RPFM (mean $= -0.006$).

4.5.2 Sources of Inefficiency

Given the encouraging results in the preceding section, further analysis of the influence of PPS on inefficiency is warranted to test for the hypotheses of Section 4.2. Table 4.3 presents estimation results for the three models outlined in Section 4.4.3. The dependent variable is the mean inefficiency $\bar{u}_{it}$ of eq. (4.2). All results are estimated together with the parameters of the cost frontier, shown in Table 4.2 for Model (1).[5]

In Model (1), an unexpected positive sign is obtained for PPS in the SFM (mean $= 0.02$), indicating that PPS increases hospital inefficiency. In contrast, the more appropriate RPFM shows a small negative value (mean $= -0.01$). This is rather counterintuitive to the findings in Figure 4.3, suggesting that unobserved heterogeneity substantively biases estimates of the influence of PPS on inefficiency. Estimates for the interaction term $PPS{:}DRG$ are more intuitive. In both cases, DRG is negatively correlated with hospi-

[5] Technology parameters of Model (2) and (3) are not shown. They are found to be comparable to those discussed in Section 4.5.1.

Table 4.3: Determinants of Inefficiency by Model Type, Years 2004-9

Variables[1)]	SFM			RPFM		
	Mean	2.5%	97.5%	Mean	2.5%	97.5%
Model 1:						
Constant	-0.13	-0.31	0.01	-0.17	-0.35	-0.04
PPS	0.02	-0.09	0.12	-0.01	-0.15	0.09
PPS:DRG	-0.15	-0.31	-0.04	-0.08	-0.22	0.02
Model 2:						
Constant	-0.15	-0.35	-0.01	-0.17	-0.33	-0.05
CASEP	-0.08	-0.30	0.13	-0.07	-0.26	0.11
GLOB	0.03	-0.07	0.14	0.00	-0.13	0.10
CASEP:DRG	-0.06	-0.29	0.14	-0.04	-0.24	0.15
GLOB:DRG	-0.15	-0.32	-0.03	-0.07	-0.22	0.03
Model 3:						
Constant	-0.16	-0.36	-0.03	-0.19	-0.36	-0.06
PPS	0.02	-0.11	0.13	-0.02	-0.16	0.09
PPS:DRG	-0.13	-0.33	0.03	-0.08	-0.25	0.07
PPS:DRG1	-0.08	-0.29	0.10	-0.03	-0.21	0.14
PPS:DRG2	-0.01	-0.25	0.18	-0.02	-0.19	0.15

[1)] Mean inefficiency $\bar{u}_{it}$ is the dependent variable in each model. Technology parameters are dropped for simplicity; those of Model (1) are shown in Table 4.2.

tal inefficiency. However, the effect is unreasonably large in the SFM (mean = -0.15). Estimates for Model (2) are similar. Although both approaches have comparable signs, estimates for $GLOB$ and $GLOB{:}DRG$ differ significantly between the two approaches. In the SFM, $GLOB$ is found to have a positive influence (mean = 0.03) on inefficiency and no effect in the RPFM (mean = 0.00). Moreover, the interaction term indicates an unreasonably high negative effect (mean = -0.15) in the SFM. Even in Model (3), which controls for a possible time-lag of DRG, the estimated effects are systematically larger in the SFM and again PPS seems to be significantly biased by unobserved heterogeneity. Taken together, estimation results are less robust between the SFM and RPFM than expected from Figure 4.3. The results mainly depend on the assumptions made to the production technology.

Nevertheless, RPS appears to undermine efforts for cost containment, addressing hypothesis (1). Inefficiency decreases by about -0.01 for hospitals with PPS, meaning that a switch to PPS causes hospitals to reduce their variable costs by an average of 1 percent. However, as shown in Model (2), efficiency gains depend substantially on whether hospitals are paid by a global budget or receive payments per patient case. While a payment per patient case system reduces hospital-specific inefficiency by about -0.07 on average, no efficiency gains occur with a global budget system (in the biased SFM they are even more inefficient, 0.03). The renegotiations that most cantons still allow of the global budget seem to reduce incentives for cost minimization.

Whereas general remuneration settings with a per diem element can be used to unnecessarily keep a patient in the hospital, a DRG system strengthens incentives for cost minimization. As estimates from Table 4.3 show, this results in 8 percent lower cost inefficiency for hospitals with a DRG system, addressing hypothesis (2). Model (2), which shows the unique effects of DRG for hospitals with a global budget and payments per patient case system, reveals that the efficiency gains of DRG are even larger in the global budget (-0.07) than in the payment per patient case system (-0.04). However, the combined effect of $CASEP$ is larger ($-0.11 = -0.07 - 0.04$) than the expected cost savings for $GLOB$ ($-0.07 = 0.00 - 0.07$), making the payment per patient case system together with DRG classifications the preferable variant for Switzerland. Under a payment per patient case system with DRG classifications, hospitals have 11 percent lower inefficiency scores on average than their counterparts with RPS. Moreover, Model (3) reveals that a DRG system is not fully effective in the first year after initiation. Although most cost savings occur in the first year ($DRG = -0.08$), additional reduction is observable in the second ($DRG_1 = -0.03$) and third year ($DRG_2 = -0.02$) after implementation.

Finally, these findings are in line with the theoretical expectations, for example outlined in Chalkley and Malcomson (2000) and Newhouse (1996). With respect to the hospital payment reform becoming effective in 2012, these results support the policy expectations that PPS will rather increase cost efficiency. However, the implementation has to be fully prospective and preclude any bailouts.

4.6 Concluding Remarks

The purpose of this article was to estimate the effectiveness of prospective payment systems in reducing hospital cost inefficiency. Hospitals in Switzerland are analyzed, which, in contrast to previous studies, enables a comparison of a retrospective per diem system with two prospective payment systems, one based on a global budget and another based on payments per patient case. Since the results of previous studies may have been affected by the existence of unobserved heterogeneity, two stochastic frontier models are used to control for potential bias. The first is a standard frontier model (SFM) that assumes a homogeneous technology for all hospitals. The second one is a random parameter frontier model (RPFM) that controls for unobserved heterogeneity with hospital group-specific intercepts and slope parameters. A variable cost frontier is estimated for approximately 90 public financed Swiss hospitals during the time period of 2004 to 2009.

There are two main results from this analysis. First, a comparison of the standard and random parameter frontier models reveals that heterogeneity is substantial between Swiss hospital categories. Inefficiency scores are biased upwards by two percent on average in the SFM. The maximum inefficiency score decreases from 0.294 in the SFM to 0.157 in the RPFM, putting the maximum cost savings at approximately 16 percent. Further analysis of the determinants of inefficiency shows that unobserved heterogeneity systematically varies among hospitals, indicating that the SFM is not able to detect the true effect of prospective payment systems on inefficiency. The assumptions made for the production technology (SFM vs. RPFM) are important in the Swiss case.

Second, prospective payment systems are associated with an increase in hospital cost efficiency, particularly for the payment per patient case system. Payment systems designed to put hospitals at operating risk seem to be more effective in reducing hospital costs than retrospective payment systems. However, these effects may be diminished if cantons do not firmly preclude a bailout. Results relating to the global budget system reveal that if hospitals can obtain higher budgets to cover past errors, then the incentive for cost minimization disappears. In addition, the settings for the remuneration per admission are also important. Whereas general remuneration settings with a per diem element can be used to unnecessarily keep a patient in the hospital, a DRG system strengthens incentives for cost minimization. Nonetheless, estimates show that DRG is not fully effective after initiation. Additional efficiency gains occur later on, although these are smaller in the third year than in the second year. Therefore, these empirical findings are in line with the theoretical expectations. With respect to the hospital payment reform effective in 2012, these results support the expectations of Swiss politicians that the new payment system can contain hospital costs. However, the implementation has to be fully prospective and has to preclude any bailouts.

This analysis is not without limitations. Above all, unobserved heterogeneity is estimated as a time-invariant random variable, meaning that all time-invariant random noise is measured as heterogeneity. Since inefficiency could be time-invariant as well, estimates to the RPFM underestimate inefficiency. Nevertheless, together with the SFM, which overestimates inefficiency, the true influence of PPS must lie somewhere between, making the results still reliable. Additionally, a translog form would have been more accurate than the Cobb-Douglas form for the production technology since it can test for specific features of technology (like economies of scale or homotheticity) by examining the estimated model parameters. Unfortunately, limitations of the data dictated the application of the reduced self-dual Cobb-Douglas form, which per definition is restricted to constant elasticities of substitution and is constant in economies of scale. Thus, estimates might be biased in cases when these assumptions are not reasonable. In spite of this limitation, the analysis not only identifies the effect of PPS on inefficiency, it also outlines the importance of unobserved heterogeneity in deriving unbiased inefficiency scores.

Appendix: The Bayesian Specification

This paper uses Bayesian statistics to estimate eqs. (4.1) and (4.3). Inference is made from a posterior distribution $p(\theta|X)$ of the unknown parameters (summarized as θ) given the observed data (summarized as X). According to the Bayesian rule this is

$$p(\theta|X) = \frac{\mathcal{L}(X|\theta)p(\theta)}{p(X)} \propto p(\theta)\mathcal{L}(X|\theta), \tag{4.9}$$

expressed as the product of the prior information $p(\theta)$ and the likelihood $\mathcal{L}(X|\theta)$, respectively.

For the estimates in Section 4.5, the posterior distribution for the SFM is specified as

$$\begin{aligned} p(\alpha, \beta, u, \gamma, \sigma_v^{-2}, \sigma_u^{-2}; C, Y, W, Z) &\propto p(\alpha, \beta, \gamma, \sigma_v^{-2}, \sigma_u^{-2}) \prod_{i=1}^{N}\prod_{t=1}^{T} p(u, \gamma, \sigma_u^{-2}|Z) \\ &\times \prod_{i=1}^{N}\prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_v^2}} exp\left[-\frac{1}{2\sigma_v^2}\left(C_{it} - [C(Y_{it}, W_{it}; \alpha, \beta) + u_{it}]\right)^2\right], \end{aligned} \tag{4.10}$$

where $p(\alpha, \beta, \gamma, \sigma_v^{-2}, \sigma_u^{-2})$ are probability distributions of the unknown parameters. The likelihood function in eq. (4.10) is as in Griffin and Steel (2007), normally distributed with σ_v^2 as the variance of the random noise $v_{it} = C_{it} - [C(Y_{it}, W_{it}; \alpha, \beta) + u_{it}]$. This is a gain in flexibility over classical maximum likelihood applications, where a joint density function of the random noise and the inefficiency term is specified. Here, only random noise enters the likelihood function. Inefficiency is estimated hierarchically as a latent variable along with the other parameters of the cost frontier.

Turning to the RPFM the posterior is given by

$$\begin{aligned} &p(\alpha, \bar{\alpha}, \beta, \bar{\beta}, u, \gamma, \Sigma, \sigma_v^{-2}, \sigma_u^{-2}; C, Y, W, Z) \propto p(\bar{\alpha}, \bar{\beta}, \gamma, \Sigma, \sigma_v^{-2}, \sigma_u^{-2}) \\ &\qquad \times \prod_{j=1}^{J} (2\pi)^{-K/2}|\Sigma|^{-1/2} exp\left[-\frac{1}{2}\left(\begin{pmatrix}\alpha_j \\ \beta_j\end{pmatrix} - \begin{pmatrix}\bar{\alpha} \\ \bar{\beta}\end{pmatrix}\right)' \Sigma^{-1} \left(\begin{pmatrix}\alpha_j \\ \beta_j\end{pmatrix} - \begin{pmatrix}\bar{\alpha} \\ \bar{\beta}\end{pmatrix}\right)\right] \\ &\qquad \times \prod_{i=1}^{N}\prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_v^2}} exp\left[-\frac{1}{2\sigma_v^2}\left(C_{it} - [C(Y_{it}, W_{it}; \alpha_j, \beta_j) + u_{it}]\right)^2\right] \\ &\qquad \times \prod_{i=1}^{N}\prod_{t=1}^{T} p(u, \gamma, \sigma_u^{-2}|Z). \end{aligned} \tag{4.11}$$

Again, the likelihood function is specified as a normal distribution and inefficiency is estimated as a latent variable together with the other unknown parameters. Different is the specification of the random intercept α_j and the slope parameters β_j, which are estimated at two levels. At the first level, overall influences on hospital costs ($\bar{\alpha}$, $\bar{\beta}$) are determined, corresponding to the first factor following the proportionally sign of eq. (4.11). The second-level estimates of the individual effects (α_j, β_j) defined in eq. (4.4) are derived from the multivariate normal distribution shown in eq. (4.11).

In contrast to classical statistics, an application of Bayesian statistics requires additional information for the prior distributions of the unknown parameters, since all parameters are considered as random variables. They should comprise all information available before any data are involved in the statistical analysis. In this case, the values for the hyperparameters are chosen in a way to imply relatively vague but proper priors. In particular, the priors for the SFM and RPFM are assumed to be independent,

$$p(\alpha, \beta, \gamma, \sigma_v^{-2}, \sigma_u^{-2}) = p(\alpha)p(\beta)p(\gamma)p(\sigma_v^{-2})p(\sigma_u^{-2}) \tag{4.12}$$

$$p(\bar{\alpha}, \bar{\beta}, \gamma, \Sigma, \sigma_v^{-2}, \sigma_u^{-2}) = p(\bar{\alpha})p(\bar{\beta})p(\gamma)p(\Sigma)p(\sigma_v^{-2})p(\sigma_u^{-2}) \tag{4.13}$$

Here, $p(\alpha) = f_N[0, \theta_\alpha]$, $p(\bar{\alpha}) = f_N[0, \theta_{\bar{\alpha}}]$, $p(\beta) = f_N[0, \theta_\beta]$, $p(\bar{\beta}) = f_N[0, \theta_{\bar{\beta}}]$ have a normal distribution with mean zero and a diffuse prior for their corresponding precision θ. The precision of the likelihood function has a gamma distribution $p(\sigma_v^{-2}) = f_G[\mu, \theta_{\sigma_v^{-2}}]$ with diffuse shape and scale parameters. Inefficiency is assumed to be truncated normally distributed $p(u, \gamma, \sigma_u^{-2}|Z) = f_N^+[\gamma Z, \sigma_u^{-2}]$ with $\sigma_u^{-2} = f_G[5, (5 * log(\bar{r})^2)]$ and $p(\gamma) = f_N[0, \theta_\gamma]/\sqrt{f_G[5, (5 * log(\bar{r})^2)]}$. This specification is in line with Griffin and Steel (2007) and Koop et al. (1997), permitting to impose a priori information with regard to mean efficiency, $\overline{eff} = exp(-\overline{u})$. Following the formulation of Griffin and Steel (2007), $\overline{eff} = 0.875$ is assumed for prior efficiency. Finally, the precision of the random parameters is specified as a Wishart distribution $p(\Sigma) = f_W[S]$ in accordance with eq. (4.11) with diffuse prior for the covariance matrix S.

Finally, note that estimates of the unknown parameters can be derived by the marginal posteriors of eqs. (4.10) and (4.11). However, it is not always possible to compute the posteriors analytically. Therefore, iterative Monte Carlo Markov Chain (MCMC) simulation is used, which involves iterative sampling from posterior parameter densities. Here, we use WINBUGS to derive the estimates (see Ntzoufras, 2009 for an introduction). The corresponding computational codes for the SFM and RPFM are shown in Table 4.4.

Table 4.4: Computation Codes for the Standard Frontier and the Random Parameter Frontier Model

Standard Frontier Model	**Random Parameter Frontier Model**

model{
for (it in 1:NT){
firm[it] ← n[it,1]
	model{
	for(it in 1:NT){
	firm[it] ← n[it,1]
	typ[it] ← n[it,3]

Likelihood: Y[it] ∼ dnorm(mu[it], prec) mu[it] ← inprod(b[1:K+1], X[it, 1:K+1]) + u[it] }	**Likelihood:** Y[it] ∼ dnorm(mu[it], prec) mu[it] ← inprod(b[typ[it],1:K+1], X[it,1:K+1]) + u[it] }

Priors:
for (it in 1:NT) {
u[it] ∼ djl.dnorm.trunc(mu1[it],lambda,0,1000)
mu1[it] ← inprod(t[1:L+1], Z[it,1:L+1])
}
	Priors:
	for (it in 1:NT) {
	u[it] ∼ djl.dnorm.trunc(mu1[it],lambda,0,1000)
	mu1[it] ← inprod(t[1:L+1], Z[it,1:L+1])
	}
	for(j in 1:J){
	for(k in 1:K+1){
	b[j,k]← xi.b[j]*b.raw[j,k]
	}
	b.raw[j,1:K+1] ∼ dmnorm(b.bar.raw[],b.tau.raw[,])
	}

for (k in 1:K+1) {
b[k] ∼ dnorm(0, 0.0001)
}
	for(k in 1:K+1){
	b.bar[k] ← xi.b[k]*b.bar.raw[k]
	b.bar.raw[k] ∼ dnorm(0,0.0001)
	xi.b[k] ∼ dunif(-10,10)
	}

for(l in 1:L+1){
t[l] ← gamma[l] / sqrt(lambda)
gamma[l] ∼ dnorm(0.0, 1)
}
lambda0 ← 5*log($\bar{r}$)*log($\bar{r}$)
lambda ∼ dgamma(5,lambda0)
prec ∼ dgamma(0.1,0.01)
	for(l in 1:L+1){
	t[l] ← gamma[l] / sqrt(lambda)
	gamma[l] ∼ dnorm(0.0, 1)
	}
	lambda0 ← 5 * log($\bar{r}$) * log($\bar{r}$)
	lambda ∼ dgamma(5,lambda0)
	prec ∼ dgamma(0.1,0.01)
	b.tau.raw[1:K+1,1:K+1] ∼ dwish(S[1:K+1, 1:K+1], nu)
	nu ← K+1
	Sigma.B.raw[1:K+1,1:K+1] ← Inverse (b.tau.raw[,])
	for(k in 1:K+1){
	Sigma.B[k] ← abs(xi.b[k])* sqrt(Sigma.B.raw[,])
	}

CHAPTER V

Efficient Provision of Electricity for the United States and Switzerland

Boris Krey, Philippe K. Widmer,
and Peter Zweifel

Abstract: This study applies financial portfolio theory to determine efficient frontiers in the provision of electricity for the United States and Switzerland. Expected returns are defined by the rate of productivity increase of power generation (adjusted for external costs), volatility, by its standard deviation. Since unobserved productivity shocks are found to be correlated, Seemingly Unrelated Regression Estimation (SURE) is used to filter out the systematic component of the covariance matrix of the productivity changes. Results suggest that as of 2003, the feasible maximum expected return (MER) electricity portfolio for the United States contains more *Coal*, *Nuclear*, and *Wind* than actual but markedly less *Gas* and *Oil*. The minimum variance (MV) portfolio contains markedly more *Oil*, again more *Coal*, *Nuclear*, and *Wind* but almost no *Gas*. Regardless of the choice between MER and MV, U.S. utilities are found to lie substantially inside the efficient frontier. This is even more true of their Swiss counterparts, likely due to continuing regulation of electricity markets.[6]

Keywords: efficiency frontier, energy, electricity, portfolio theory, Seemingly Unrelated Regression Estimation (SURE)

JEL classification: C32; G11; Q49

6 This research was supported by the Swiss Federal Office of Energy under the supervision of CORE, the Federal Energy Research Commission. The authors would like to thank Andreas Gut, Matthias Gysler, Lukas Gutzwiller, Tony Kaiser, Michel Piot, and Pascal Previdoli as well as the participants in the annual SSES meetings (Lugano, March 2006 and Zurich, March 2005), IAEE conferences (Florence, June 2007, Potsdam, June 2006 and Taipei, June 2005) and the Infrastructure Days (Berlin, October 2006 and October 2005) for many helpful comments. Shimon Awerbuch also provided valuable suggestions. Remaining errors are our own.

5

Efficient Provision of Electricity for the United States and Switzerland

5.1 Introduction

Like most industrial countries, the United States and Switzerland face great challenges in the provision of electricity arising from dwindling domestic resources. Both countries are expected to confront substantial shortfalls during the next twenty years. According to the U.S. National Energy Policy Development Group (NEPG), the projected gap amounts to nearly 50 percent of 2020 demand. As for Switzerland, a study conducted by the Paul Scherrer Institute (PSI) estimates a shortfall of 20 percent by 2020 (Gantner et al., 2000).

The solutions available to the two countries are the same, viz. import more power (from Canada and France, respectively) or increase domestic supply by investing in new generating technologies. Especially with the latter strategy, there is a substantial interest in providing electricity as economically as possible. Therefore, the question of this paper is, can the United States and Switzerland improve efficiency in their provision of electricity? If so, what are the attractive technologies, taking into account external costs that sooner or later will be factored into electricity prices?

For the measurement of efficiency, Stochastic Frontier Analysis (SFA) and Data Envelopment Analysis (DEA) are the two dominant alternatives. Fundamentally, both approaches assume a uniform production technology to infer the efficient use of a technology from observed choices of input and output quantities by firms. However, these approaches only work well when productive units are homogenous with regard to technology and face stable input prices and hence little uncertainty (see e.g. Greene, 2004a). In the provision of electricity, these circumstances are not satisfied for at least two reasons.

1. *Heterogenous technologies:* Each power plant has its own type of technology, depending on its primary energy source ($Coal$, Gas, $Hydro$, $Nuclear$, Oil, $Wind$). The issue therefore is not the cost-minimizing use of one common technology but determining an optimal portfolio of electricity-generating technologies.

2. *Cost uncertainty:* Exogenous shocks (e.g. the Gulf war in the case of oil) cause unexpected changes in input prices which affect the level and development of unit cost.

Therefore, it is not sufficient to merely focus on least-cost provision of electricity; in view of a portfolio of technologies with uncertain cost characteristics, the optimal mix of technology becomes the issue.

Such a mix can be determined by applying mean-variance portfolio theory (see e.g. Kienzle and Andersson, 2008, Krey, 2008, Awerbuch, 2006, Yu, 2003, Berger et al., 2003, Humphreys and McClain, 1998, and Bar-Lev and Katz, 1976). Here, a social planner (e.g. federal government) is assumed to act like a financial investor, who hedges against the ups and downs of the market by holding a diversified portfolio of securities. In contrast to a least-cost strategy, capacity planning does not only reflect productivity but also risk at a given level of productivity. Indeed, the objectives of the U.S. NEPG support the portfolio approach to electricity advocated here (see NEPG, 2004). The methodological innovation of this paper consists in recognizing that there are common shocks impinging on the production frontiers and hence the development of productivity[1] in generating technologies. Taking this correlation into account in the estimation of the covariance matrix (using Seemingly Unrelated Regression Estimation, SURE) can give rise to important gains in

[1] From an investor perspective, rates of return associated with a particular generating technology are decisive. However, the available data do not permit to track price-cost margins. As a proxy, relative productivity changes will be used, which reflect relative unit cost changes (but have the advantage of being positively defined in the case of development.)

the efficiency of estimation. To the best of the authors' knowledge, SURE has not been applied yet to the calculation of efficient electricity portfolios.

A comparison between the United States and Switzerland is of interest for several reasons. First, both countries heavily rely on imported fuels (*Gas* and *Nuclear*, respectively) for their power generation. While they can purchase primary energy sources at market prices, there are differences in their technology mix, giving rise to the question of whether this reflects differences in efficiency. Second, insights may be expected with regard to regulation. Contrary to the United States, the Swiss electricity market continues to be highly regulated. Swiss voters rejected liberalization efforts in a popular referendum at the end of 2002 (see EMG, 2000 and EMV, 2002). The usual presumption would be that U.S. power generation is closer to the efficient frontier than its Swiss counterpart. Finally, several countries (notably China and India) have to meet a rapidly increasing demand for electricity. For them, it is of considerable importance to invest in energy sources in a way that avoids inefficiency. This contribution should provide some help towards achieving that objective.

Results show that returns and volatilities differ greatly between technologies and between SURE and OLS estimates. While optimal choice depends on risk aversion (which is not known), the maximum expected return (MER) and the minimum variance (MV) portfolios constitute two extreme solutions. The feasible MER portfolio for the United States contains more *Coal*, *Nuclear*, and *Wind* than actual but markedly less *Gas* and *Oil*; the MV portfolio combines more *Oil*, *Coal*, *Nuclear*, and *Wind* but almost no *Gas*. Regardless of the choice between MER and MV, U.S. utilities are found to lie substantially inside the efficient frontier. This is even more true of their Swiss counterparts, likely due to continuing regulation of electricity markets.

This paper is structured as follows. Section 5.2 contains a short description and critique of the conventional least-cost planning approach. Sectoral optimization is shown to result in inefficiency in the presence of productivity shocks. However, the proposed mean-variance portfolio approach requires a stable variance-covariance matrix of returns. The construction of this matrix based on SURE is explained in Section 5.3. Section 5.4 is devoted to the empirical application to U.S. and Swiss data. First, the data base and the SURE and OLS specifications are described. Econometric results ar presented and then used in the determination of efficient mean-variance frontiers with and without constraints imposed. Section 5.5 contains a summary and suggestion for future research.

5.2 From Least-Cost Planning to Optimal Provision of Electricity

This section expands the relationship between least-cost planning of electricity supply and overall optimal provision of electricity. Traditionally, research has focused on identifying power plants using a particular technology (e.g. gas as fuel) that achieve maximum productivity (see e.g. Diewart and Nakamura (1999) and Kumar and Gupta, 2004). With the advent of deregulation of power generation in the United States and the European Union, this type of research has been concentrating on the distribution sector (see e.g. Resende, 2002 and Farsi et al., 2008). However, the idea continues to be to allocate output to the most productive (or least-cost, respectively) units. This sectoral approach rests on the following concepts. Let there be a production process $Y = f(X)$, mapping input quantities $X = (x_1, ..., x_m)$, $X \in \mathbb{R}^m_+$ into s output quantities $Y = (y_1, ..., y_s)$, $Y \in \mathbb{R}^s_+$. The *production set* is defined by (Koopmans, 1951 and Debreu, 1951)

$$\Gamma_j = \{(X_j, Y_j)|\ Y_j \leq f(X_j)\}, \tag{5.1}$$

describing all possible combinations (X_j, Y_j). For illustration purposes, Figure 5.1A shows the production set Γ_j for a single input (generating costs) and single output (kilowatt-hours produced) for gas-fueled power plants. The combinations of interest are those on the boundary of Γ_j which are technically efficient, meaning that for a given quantity of input $\bar{X}_j$ no more output Y_j can be produced or inversely, no less input X_j can produce a given output $\bar{Y}_j$. According to Shephard (1970) the boundary can be expressed as an input or an output isoquant

$$\begin{aligned} Iso\ X(y) &= \{x|\ x \in X(y), \theta x \notin X(y), \forall\ 0 < \theta < 1\} \\ Iso\ Y(x) &= \{y|\ y \in Y(x), \theta^{-1} y \notin Y(x), \forall\ 0 < \theta < 1\}, \end{aligned} \tag{5.2}$$

with θ denoting a scalar by which all inputs can be reduced without leaving the feasibility set or becoming technically efficient, respectively. Accordingly, θ^{-1} symbolizes the scaling-up factor for the outputs.

Furthermore, overall productivity of domestic supply can be increased by investing in those technologies $\Psi_{least-cost}$ that are most productive, satisfying

$$\Psi_{least-cost} = \arg\max\{\Gamma | \Gamma_j \in \Gamma\}. \tag{5.3}$$

However, as already argued in the introduction, in power generation technology is not stable over time. In Figure 5.1A, the production set Γ_j ($j = Gas$) moves down between periods t and $t+1$ due to a negative productivity shock. In this case, a least-cost strategy may be inappropriate for domestic supply. Let G_t be one of the efficient gas-based power companies. If it is to maintain its contribution to electricity supply ($\bar{Y_j}$), it would have to use much more *Gas* (as indicated by point G'_{t+1}), imparting a cost shock to total supply. To the extent that other technologies (e.g. *Hydro*) are not affected by the shock, a reallocation in favor of these technologies is indicated. In the extreme, this would amount to holding the company to its initial input (and cost) level, causing it to move to point G''_{t+1}. The associated shortfall in power supply would have to be made up by companies using other technologies, causing them to deviate from their least-cost allocations.

A possible way to overcome the problem of inefficiency due to stochastic shocks is to account for technology risk, e.g. indicated by the variance of efficient frontier determined by the input or output isoquant. Deviations from least-cost planning now become possible if the technology considered differs from the others in terms of risk.

However, this decision rule is still sectoral, failing to benefit from the possible risk diversification effects offered by a portfolio of generation technologies. Acting like a forward-looking investor, the social planner limits his choice to the set of efficient portfolios. These are portfolios that for a given level of risk $\bar{\sigma_p^2}$ offer the highest expected return or conversely, for a given level of expected return $\bar{R}_p$ offer the minimum risk. This is the solution of two equivalent optimization problems (Markowitz, 1952),

$$\max_{w_j} E(R_p) \quad s.t. \quad w'1 = 1,\ w'\Sigma w \leq \bar{\sigma_p^2}, \tag{5.4}$$

$$\min_{w_j} \sigma_p^2 \quad s.t. \quad w'1 = 1, w'E(R) \geq \bar{R}, \tag{5.5}$$

with w as the vector of weights and the expected return $E(R_p)$ given by

$$E(R_p) = [w_1 \cdots w_J] \begin{bmatrix} E(R_1) \\ \vdots \\ E(R_J) \end{bmatrix} = w'E(R), \quad \text{with} \quad \sum_{j=1}^{J} w_j = 1. \tag{5.6}$$

The volatility of the portfolio's expected return involves not only the respective variances but all the covariances as well. Therefore, one has for the variance σ_p^2,

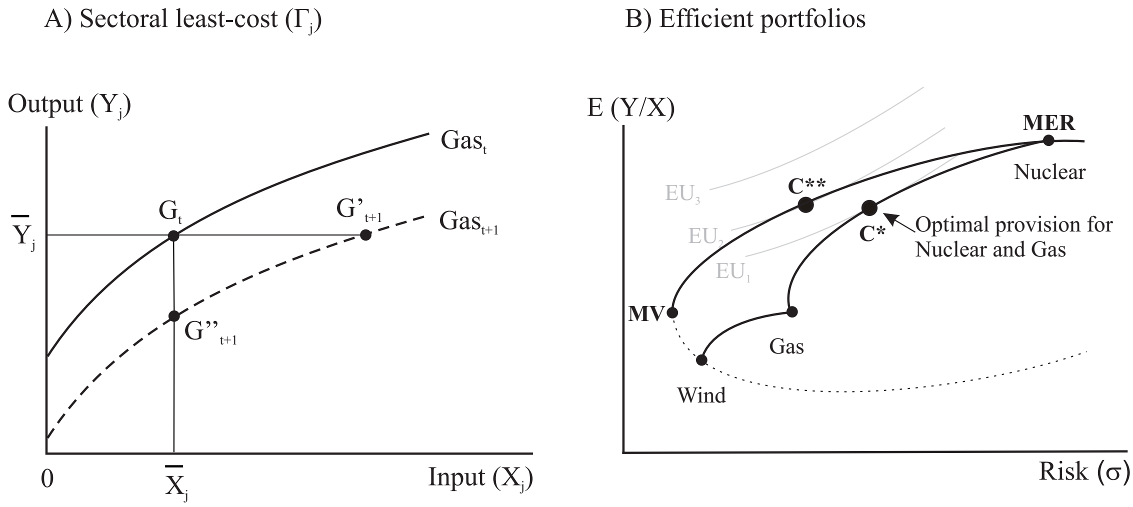
$$\sigma_p^2 = [w_1 \cdots w_J] \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1J} \\ \vdots & \ddots & \vdots \\ \sigma_{J1} & \cdots & \sigma_{JJ} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_J \end{bmatrix} = w'\Sigma w, \quad \text{with} \ \ w'\Sigma w > 0. \tag{5.7}$$

In both formulations (5.4) and (5.5), the decision variables are the weights w_j assigned to the components of the portfolio, i.e. the generating technologies in the present context.

Figure 5.1B illustrates the case of three generating technologies *Wind*, *Gas*, and *Nuclear*. By assumption, *Wind* and *Gas* have a low risk with low expected return, while *Nuclear* has a high risk that is negatively correlated to *Wind* and *Gas*. The horizontal axis depicts risk, the vertical axis displays the expected return, respectively. For the moment, this is taken to be expected productivity E(Y/X); see Section 5.4.1 for more detail. Three possible values of Y_j/X_j can be read off Figure 5.1A as the slopes of rays from the origin through points G_t, G'_{t+1}, and G''_{t+1}, respectively. Note that they are defined by the least-cost productive units for a given technology.

Starting with *Gas* and *Nuclear*, the social planner can allocate domestic supply between these two risky technologies. Without incorporating risk at all, least-cost planning would call for a complete reliance on *Nuclear* because it offers the maximum expected

Figure 5.1: Sectoral Least-cost and Portfolio Efficient Frontiers (Electricity)



return (MER). A sectoral approach incorporating variances but neglecting the fact that shocks can be less than perfectly correlated would result in a linear combination of *Gas* and *Nuclear*, with weights w_j inversely proportional to the respective variances (Zweifel and Eisen, 2011, ch. 4.1.2). The solution of the optimization problem (4) or (5), respectively results in the semi-elliptic efficient frontier linking *Gas* and *Nuclear*. The mix of the two technologies varies along this frontier. The lower the coefficient of correlation between shocks affecting *Gas* and *Nuclear*, the more marked is the concavity of the frontier, indicating benefits of risk diversification. However, choice of the optimal portfolio depends on the degree of risk aversion of the investor, reflected by the slope of his or her indifference curves (marked EU for constant expected utility). As long as there are only *Gas* and *Nuclear* and given moderate risk aversion, the optimum is given by point C^*, the point of tangency of the efficient frontier and the indifference curve.

Now let there be a third technology (*Wind*). This creates additional opportunities for diversification, shifting the efficient frontier upward and inward. As before, knowledge of the investor's risk preference would be necessary to predict the choice of optimal provision (C^{**}). But because this knowledge is not available (at least concerning the provision with electricity) for the United States and Switzerland, two extreme solutions are worth pointing out. As can be gleaned from Figure 5.1B, a very risk-averse investor is predicted to opt for the minimum variance (MV) provision. By way of contrast, an (almost) risk-neutral investor prefers the maximum expected return (MER) alternative, usually implying a very different mix of generating technologies. Comparing these two extreme solutions permits to assess the maximum influence of risk aversion on the optimal provision of electricity.

Note that the portfolio approach does not revolve around single technologies, but an efficient mix of several technologies. Even if a particular technology is dominated by others in terms of risk and expected return, it may still contribute to the optimal provision of electricity because of its diversification effect.

5.3 Construction of a Stable Variance-Covariance Matrix (Σ)

One important condition for calculating the optimal allocation of generating technologies is the estimation of a stable variance-covariance matrix Σ. An unstable estimate of Σ would result in highly variable optimal weights w_j^* of technologies [see eq. (5.6)]. Lack of stability can be due to extreme shocks, which may cause outliers during several years.

One possibility that is widely suggested in financial literature (see e.g. Bodie et al., 2005, ch. 13) to achieve time-invariant estimates is the generalized autoregressive conditional heteroscedasticity (GARCH) or the autoregressive (AR) model of eq. (5.8),

$$R_{j,t} = \beta_{j,0} + \sum_{n=1}^{N} \beta_{j,n} * R_{j,t-n} + \varepsilon_{j,t}, \tag{5.8}$$

where $R_{j,t}$ is the return of technology j in year t, $\beta_{j,0}$ is a constant for technology j, $\beta_{j,n}$ is the coefficient pertaining to the returns lagged n years, $R_{i,t-n}$ is the dependent variable lagged n years, and $\varepsilon_{j,t}$ is the error term pertaining to technology j in year t.

However, while this formulation suffices to insulate expected conditional values $\overline{R_{j,t}}$ from extreme shocks (which would spill over into the estimated correlation matrix), Krey and Zweifel (2009) find that due to unobserved common shocks impinging on technologies at the same time, SURE (Seemingly Unrelated Regression Estimation) is the more efficient alternative. The correlation between the error terms $\varepsilon_{j,t}$ constitutes information that can be used in SURE to obtain sharper estimates of the β parameters (see Section 5.4.2 for empirical evidence).

In the example above, the SURE model consists of three regression equations (for *Wind*, *Gas*, and *Nuclear*), each of which satisfies the assumptions of the standard regression model,

$$\begin{aligned} R_{1,t} &= a_0 + \sum_{n=1}^{N} a_{1,n} R_{1,t-n} + \varepsilon_{1,t} \\ R_{2,t} &= b_0 + \sum_{n=1}^{N} b_{1,n} R_{2,t-n} + \varepsilon_{2,t} \\ R_{3,t} &= c_0 + \sum_{n=1}^{N} c_{1,n} R_{3,t-n} + \varepsilon_{3,t}, \end{aligned} \tag{5.9}$$

where $R_{1,t}$, $R_{2,t}$, $R_{3,t}$ are the returns of technologies $j = 1, 2, 3$ in year t. a_0, b_0, and c_0 are their respective constants, $a_{1,n}$, $b_{2,n}$, $c_{3,n}$ are the coefficients of returns lagged n years, $R_{1,t-n}$, $R_{2,t-n}$, $R_{3,t-n}$ are the dependent variable lagged n years, and $\varepsilon_{1,t}$, $\varepsilon_{2,t}$, $\varepsilon_{3,t}$ are the error terms with $E(\varepsilon_{j,t}) = 0$, and $E(\varepsilon_{i,t}\varepsilon_{j,s}) = \sigma_{i,j}$ if $t = s$ and $= 0$ if $t \neq s$. This is the SURE specification, admitting nonzero contemporaneous correlations between error terms. Thus, the variance-covariance matrix Σ of residuals is not diagonal,

$$\Sigma = E(\varepsilon\varepsilon') = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} & \sigma_{1,3} \\ \sigma_{2,1} & \sigma_{2,2} & \sigma_{2,3} \\ \sigma_{3,1} & \sigma_{3,2} & \sigma_{3,3} \end{bmatrix}. \tag{5.10}$$

By way of contrast, OLS estimation would be superior if the variance-covariance matrix were diagonal. However, this does not hold for U.S. and Swiss power technologies (see Section 5.4.2).

In sum, SURE allows to simultaneously estimate the expected returns and the variances for all power generation technologies in one regression, taking into account possible correlations of error terms across equations.

5.4 Empirical Application to U.S. and Swiss Electricity Data

In this section, theory and data are combined for the construction of efficient frontiers for electricity-generating technologies in the United States and Switzerland. This calls for an estimate of expected returns $E(R_j)$ for each technology j that potentially is part of the optimal provision, of its variance σ_j^2, and its covariances σ_{ij}. Estimates of these quantities come from the SURE results shown in Section 5.4.2. Results presented in Sections 5.4.3 and 5.4.4 contrast the actual portfolio (AP) of both countries with the minimum variance (MV) and maximum expected return (MER) alternatives, which correspond to the optimum allocation in case of extremely marked and very weak degrees of risk aversion, respectively (see Section 5.2).

5.4.1 Data and Model Specifications

This article uses time-series data on annual power generation returns for several technologies. Contrary to the theoretical exposition in the preceding sections, which is in terms of productivity levels for simplicity, returns are measured as annual changes in productivity, with productivity equated to kWh electricity produced per U.S. cent.[2] This definition is similar to that of Berger et al. (2003) and Awerbuch and Berger (2003), who point out that a rational investor more likely relies on productivity changes than levels when choosing the technology mix for the future. In full analogy, a financial investor buys a stock in view of its expected future change in value rather than its current price.

The U.S. data set consists of five variables; *Coal, Nuclear, Gas, Oil*, and *Wind* power[3], covering the years 1982 to 2003. Unfortunately, no data were available for *Hydro*, which contributed an estimated 6 to 10 percent to total U.S. power supply. Nevertheless, more than 90 percent of U.S. generating capacity is covered, going beyond earlier work that was

[2] The mean value of the exchange rate for the year 2000 was used to convert Swiss cents into U.S. cents, as published by the U.S. Federal Reserve (http://research.stlouisfed.org).

[3] Data for *Coal, Nuclear, Gas* and *Oil* were obtained from the UIC (2005). *Wind* (State Hawaii, USA (www.state.hi.us) and U.S. Department of Energy (www.energy.gov)). Since the *Wind* data was not available for every year, values for 1983, 1985-1987, 1989-1994, 1996-1999 were generated by cubic spline interpolation (Knott, 2000).

limited to three technologies (Awerbuch, 2006 and Humphreys and McClain, 1998). The Swiss data on *Nuclear*[4] cover the years 1986 to 2003, those on *Run of river*[5] and *Storage hydro*[6] 1993 to 2003, and *Solar*[7], 1991 to 2003. Throughout, generation costs comprise (i) fuel costs, (ii) costs of current operations, and (iii) capital user cost[8] (depreciation of book value plus interest). In the case of *Nuclear*, decommissioning and waste disposal are also included.

For both countries, an externality surcharge for environmental damage caused by power generation is added to the costs of each technology. As in previous studies (Awerbuch, 2006 and Awerbuch, 2005), these cost data are available for total production only, precluding a differentiation according to load segments. Finally, from an efficiency point of view, the price of a product should reflect external costs only to the extent that the marginal benefit of internalization effort still covers its marginal cost. However, the externality surcharge corresponds to total estimated external cost per kWh electricity, reflecting the implicit assumption that full internalization is optimal. The data on external costs were obtained from the European Commission (2003) for the United States and Hirschberg and Jakob (1999) for Switzerland.[9] While external costs related to health and global warming do enter calculations, no data are available for some other categories such as external costs related to agriculture and forestry. In this paper, the upper bound of social cost estimates is adopted for both countries.

The resulting productivity levels are displayed in Table 5.1 together with the shares of technologies in the U.S. and Swiss power generation portfolios. As noted in panel A, the U.S. mix predominantly consists of fossil fuels (56 percent *Coal*, 18 percent *Gas*, and 3 percent *Oil*), with *Nuclear* accounting for another 21 percent of production. However, with externality surcharges included, these weights do not reflect productivity levels. While the productivity of *Coal* with its share of 56 percent is some 11 kWh per U.S. $

4 Data sources: KKl (2005), KKG (2005)
5 Data source: personal correspondence
6 Data source: personal correspondence
7 RWE (2005); The average exchange rate of 2000 was used to convert Euro cents into U.S. cents (source: U.S. Federal Reserve). RWE data from Germany is used as a proxy for Swiss solar electricity data, since solar generation technologies in both countries are similar.
8 Capital user cost can be defined in several ways. The variant "linear depreciation and interest" is used here exclusively due to lack of source data that would permit to calculate other variants.
9 No external cost data for the United States were available; therefore data from the United Kingdom were used (European Commission, 2003).

(busbar) in 2003, *Wind* power is far more productive but accounts for 2 percent only. By way of contrast, the Swiss portfolio in panel B seems to match productivity levels much better. It relies heavily on highly productive *Hydro* (32 percent *Storage hydro*, 24 percent *Run of river*); *Nuclear* accounts for 40 percent, *Solar* (a proxy of all renewable and conventional thermic technologies with a low productivity of 2 kWh per U.S. $), for a mere 4 percent.

Table 5.1: Current Portfolio Weights (Percent) and Productivity Levels (kWh/U.S. $, Prices of 2000)

Panel A: United States*					Panel B: Switzerland				
Technology	Weights		Productivity		Technology	Weights		Productivity	
	1995	2003	1995	2003		1995	2003	1995	2003
Nuclear	21	21	17	26	*Nuclear*	39	40	20	29
Coal	57	56	9	11	*Storage hydro*	27	32	39	52
Gas	17	18	16	13	*Run of river*	32	24	18	25
Oil	3	3	9	10	*Solar*	2	4	1	2
Wind	2	2	18	23					

* Excluding hydro
Source: Swiss Federal Office of Energy (SFOE) (2003), IAE (2005)

However, recall that productivity changes rather than levels are relevant for investors, who would have wanted to buy into Swiss *Solar* in 1995 in view of its rapid productivity increase in the course of nine years. From an investor point of view, Swiss *Solar* should therefore figure prominently in an efficient portfolio unless it has extremely unfavorable diversification properties.

Finally, the SURE models need to be specified. Eqs. (5.11) display the U.S. specifications that have the best statistical properties (see Section 5.4.2 below), selecting 2003 as the year of reference for the efficient portfolios,

$$\begin{aligned}
R_{Nucl,03} &= n_0 + n_1 R_{Nucl,02} + n_2 Trend_t + \varepsilon_{Nucl,03} \\
R_{Coal,03} &= c_0 + c_1 R_{Coal,02} + c_2 Trend_t + \varepsilon_{Coal,03} \\
R_{Gas,03} &= g_0 + g_1 R_{Gas,02} + g_2 R_{Gas,01} + g_3 R_{Gas,00} + g_4 Trend_t + \varepsilon_{Gas,03} \\
R_{Oil,03} &= b_0 + b_1 R_{Gas,02} + b_2 R_{Gas,01} + b_3 R_{Gas,00} + b_4 R_{Gas,99} + b_5 Trend_t + \varepsilon_{Oil,03} \\
R_{Wind,03} &= d_0 + d_1 R_{Wind,02} + d_2 Trend_t + \varepsilon_{Wind,03}.
\end{aligned} \tag{5.11}$$

Generally, influences such as technological change are hypothesized to influence productivity of electricity generation and hence returns. However, estimating such a comprehensive model is beyond the scope of this study. Rather, the relative productivity change of *Nuclear* in the United States in the year 2003 e.g., $R_{Nucl,03}$, is related to a constant (n_0), the productivity changes in the preceding year $R_{Nucl,02}$, and a time trend ($Trend_t$).

In analogy, the productivity change of *Nuclear* in Switzerland in the year 2003, $R_{Nucl,03}$, is related to a constant (n'_0), the productivity changes in the preceding years $R_{Nucl,02}$, $R_{Nucl,01}$, $R_{Nucl,00}$, and $R_{Nucl,99}$, and a time trend ($Trend_t$). The remaining eqs. (5.12) refer to *Run of river (Ror)*, *Storage hydro (Sh)*, and *Solar* (*Solar* also includes other renewable energy sources such as waste),

$$\begin{aligned}
R_{Nucl,03} &= n'_0 + n'_1 R_{Nucl,02} + n'_2 R_{Nucl,01} + n'_3 R_{Nucl,00} + n'_4 R_{Nucl,99} + n'_5 Trend_t + \varepsilon'_{Nucl,03} \\
R_{Ror,03} &= r'_0 + r'_1 R_{Ror,02} + r'_2 Trend_t + \varepsilon'_{Ror,03} \\
R_{Sh,03} &= h'_0 + h'_1 R_{Sh,02} + h'_2 Trend_t + \varepsilon'_{Sh,03} \\
R_{Solar,03} &= s'_0 + s'_1 R_{Solar,02} + s'_2 R_{Solar,01} + s'_3 R_{Solar,00} + s'_4 R_{Solar,99} + s'_5 Trend_t + \varepsilon'_{Solar,03}. \qquad (5.12)
\end{aligned}$$

5.4.2 Preliminary Testing and Econometric Results

The objective is to obtain a stable estimate of the covariance matrix Σ derived from eqs. (5.11) and (5.12). In order to be able to filter out the systematic (trend stable) component of Σ, changes in productivity must form stationary time series. Given non-stationarity, the estimate of Σ would shift over time, precluding the estimation of a reasonably stable efficient frontier [Wooldridge (2003), ch. 11]. To test for stationarity, the augmented Dickey-Fuller (ADF) test was applied. Results indicate at the one percent significance level that all productivity changes in the U.S. and Swiss data sets are stationary. To determine the correct lag order for the SURE regressions, several tests were applied, viz. Akaike's information criterion (AIC), Hannan & Quinn's information criterion (HQIC), Schwartz's Bayesian information criterion (SBIC), and the likelihood ratio test (LR) (Al-Subaihi, 2002 and Liew, 2004). The results for the U.S. data suggest five lags for *Oil*, three lags for *Gas*, and one lag for *Coal*. One lag was used for *Wind* and *Nuclear*, based on considerations of goodness of fit in SURE. The results for the Swiss data suggest four lags for *Nuclear*, while in the case of *Storage hydro* and *Run of river*, one lag suffices. Tests are inconclusive for *Solar*. However, Liew (2004) shows that lag selection tests may lack validity if the sample is small. Using a sample size of 25 he finds that the probability of correctly estimating the true order of an autoregressive process ranges between 58 percent

(SBIC) and 60 percent (HQIC). In view of the inconclusive evidence and the fact that the coefficients on the autoregressive variables used in the SURE procedure are significant without exception, four lags were applied throughout in the case of Switzerland for *Solar*.

Table 5.2: Econometric Results, United States (1982-2003)

Panel A: Results of SURE regression, dependent variable R_t: relative productivity change

	$\bar{R}$	St.D.	Const.	R_{t-1}	R_{t-2}	R_{t-3}	R_{t-4}	R_{t-5}	Trend	Obs	R^2
Coal	5.2	2.0	−0.09**	0.02					0.003**	17	0.67
Nuclear	5.8	1.8	−0.05*	0.38*					0.001	17	0.07
Gas	3.9	11.7	−0.32**	0.10	-0.89**	0.12			0.018**	17	0.67
Oil	2.5	10.4	−1.05**	−0.96**	-1.35**	−1.17**	-1.21**	-0.62*	0.050**	17	0.67
Wind	5.4	6.9	−0.03	0.73**					0.001	17	0.51

Panel B: Results of OLS regression, dependent variable R_t: relative productivity change

	$\bar{R}$	St.D.	Const.	R_{t-1}	R_{t-2}	R_{t-3}	R_{t-4}	R_{t-5}	Trend	Obs	R^2
Coal	4.8	1.5	−0.06**	0.22**					0.002*	21	0.36
Nuclear	4.8	2.3	−0.01	0.30					−0.002	21	0.21
Gas	3.6	10.5	−0.26*	0.13	-0.78**	0.23			0.015*	19	0.69
Oil	2.5	9.7	−0.91*	−0.85*	-1.21**	−0.94	-1.10*	-0.43	0.043*	17	0.62
Wind	4.1	2.6	−0.05*	0.21*					0.002	21	0.72

** significant at 1 percent level, * significant at 5 percent level.

The resulting SURE and OLS regressions are displayed in Table 5.2 for the United States and Table 5.3 for Switzerland together with their estimated average returns $\bar{R}$ and standard deviations $St.D.$ Comparing the results of SURE and OLS estimates, the first thing to note is that due to its fuller use of information, SURE results in sharper coefficient estimates than OLS. In the regressions for the United States, 15 SURE but only 13 OLS coefficients out of a theoretical total of 35 are significant at the 5 percent level or better. In the regression for Switzerland, 14 SURE but only 1 OLS coefficient out of a theoretical total of 24 are significant. This difference is of importance because the objective is to filter out transitory shocks in productivity development for obtaining a stable estimate of the variance-covariance matrix Σ. Clearly, SURE estimates serve this purpose better than their OLS counterparts. Also, the contrasts between estimates are sometimes striking. Notably, the SURE results of Table 5.2 (col. “Const.”, panel A) suggest a productivity-decreasing drift of 5 percent p.a. in American *Nuclear*, while according to the OLS estimate (panel B), the hypothesis of no drift cannot be rejected. In the case of *Wind*, it is the other way round. In the Swiss regressions, *Solar* exhibits the expected upward productivity shift in the SURE estimation (panel A of Table 5.3), which would have not been recognized as significant in the OLS alternative (panel B).

Table 5.3: Econometric Results, Switzerland (1986-2003)

Panel A: Results of SURE regression, dependent variable R_t: relative productivity change

	$\bar{R}$	St.D.	Const.	R_{t-1}	R_{t-2}	R_{t-3}	R_{t-4}	Trend	Obs	R^2
Nuclear	−3.6	12.9	−0.04	0.74**	0.93**	1.22**	1.37**	0.18**	9	0.74
Run of river	−4.1	18.6	−0.33	0.70**				0.20	9	0.51
Storage hydro	−1.2	12.0	−0.25	0.72**				0.02	9	0.22
Solar	6.7	1.0	0.34**	0.73**	0.56*	0.61*	0.55*	−0.01**	9	0.63

Panel B: Results of OLS regression, dependent variable R_t: relative productivity change

	$\bar{R}$	St.D.	Const.	R_{t-1}	R_{t-2}	R_{t-3}	R_{t-4}	Trend	Obs	R^2
Nuclear	4.3	2.2	0.10	0.03	0.29	0.14	0.38	−0.001	14	0.38
Run of river	−1.6	1.6	−0.11	0.64*				0.01	10	0.44
Storage hydro	−0.8	9.1	−0.20	0.54				0.01	10	0.35
Solar	6.7	1.0	0.32	0.69	0.60	0.58	0.40	−0.01	9	0.64

** significant at 1 percent level, * significant at 5 percent level.

The presence of correlations across equations is of interest because they determine the diversification effects in the portfolio. Panel A of Table 5.4 does indicate some negative correlations in the SURE residuals for the United States, with that between *Wind* and *Coal* attaining a value of −0.4246. Panel B of Table 5.4 permits a comparison with OLS residuals. While the estimated correlation coefficient for *Wind* and *Coal* would have been similar with −0.4062, correlation coefficients between *Nuclear* and *Coal* are less marked than their SURE counterparts. A striking difference can be seen in the case of *Gas* and *Wind*. The correlation between the SURE residuals is positive, while that between OLS residuals is negative.

In the case of Switzerland (Table 5.5), the highest partial correlation coefficient between SURE residuals (panel A) is obtained for *Solar* and *Nuclear* (0.5933), followed by *Run of river* and *Storage hydro* (0.5054). In the latter case, the common unobserved shock clearly is weather conditions, in particular the amount of precipitation. The pertinent correlation coefficients between OLS residuals (panel B) are somewhat larger with

Table 5.4: Correlation Matrices for the United States

Panel A: Partial correlation coefficients for $\widehat{\varepsilon_{i,t}}$ residuals from eqs. (5.11), using SURE (1982-2003)

	Coal	*Nuclear*	*Gas*	*Oil*	*Wind*
Coal	1				
Nuclear	−0.1140	1			
Gas	0.7605	0.0113	1		
Oil	−0.3317	0.4461	−0.2621	1	
Wind	−0.4246	−0.2520	0.1150	−0.1492	1

Panel B: Partial correlation coefficients for $\widehat{\varepsilon_{i,t}}$ residuals from eqs. (5.11), using OLS (1982-2003)

	Coal	*Nuclear*	*Gas*	*Oil*	*Wind*
Coal	1				
Nuclear	−0.0329	1			
Gas	0.7050	−0.0004	1		
Oil	−0.2835	0.3670	−0.1362	1	
Wind	−0.4062	−0.1644	−0.2073	0.0998	1

0.7201 for *Solar* and *Nuclear* and about the same for *Run of river* and *Storage hydro* with 0.5066.

Table 5.5: Correlation Matrices for Switzerland

Panel A: Partial correlation coefficients for $\widehat{\varepsilon_{i,t}}$ residuals from eqs. (5.12), using SURE (1986-2003)

	Nuclear	*Storage hydro*	*Run of river*	*Solar*
Nuclear	1			
Storage hydro	−0.4644	1		
Run of river	−0.2685	0.5054	1	
Solar	0.5933	0.0367	−0.5907	1

Panel B: Partial correlation coefficients for $\widehat{\varepsilon_{i,t}}$ residuals from eqs. (5.12), using OLS (1986-2003)

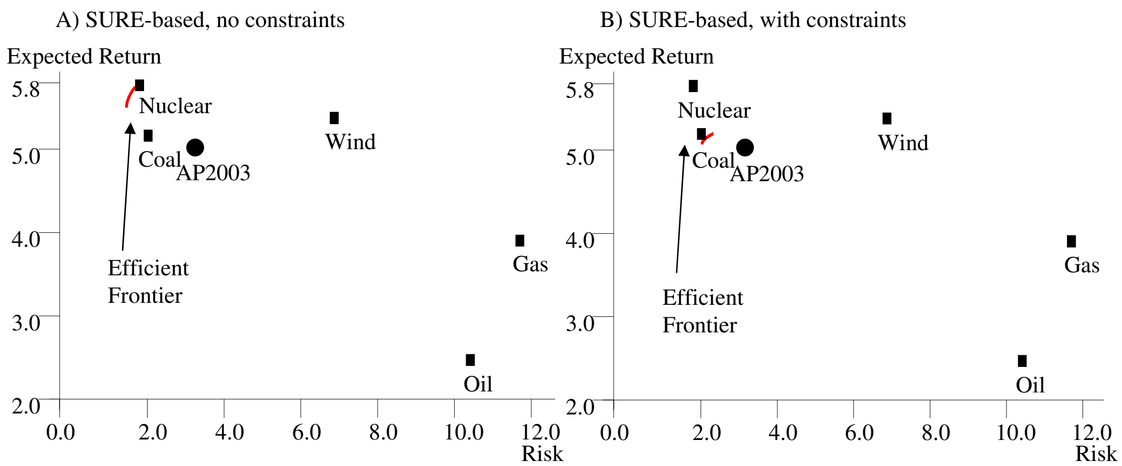
	Nuclear	*Storage hydro*	*Run of river*	*Solar*
Nuclear	1			
Storage hydro	0.3111	1		
Run of river	−0.0550	0.5066	1	
Solar	0.7201	0.2056	−0.3824	1

In sum, in contradistinction to previous studies which solely relied on OLS estimates, SURE is found to benefit from substantial correlations between unobserved shocks impinging on generating technologies. Therefore, the pertinent estimate of the variance-covariance matrix of returns Σ can be expected to be more stable than its OLS-based counterpart.

5.4.3 Efficient Provision of U.S. Electricity

Figure 5.2A displays the efficient frontier for the provision of electricity in the United States, along with the actual portfolio (AP) of 2003. No constraints are imposed on the optimization problem [see eqs. (5.4) and (5.5)] at this time. If the country's sole interest were to maximize expected return (thus maximizing the expected decrease of power generation costs), it would choose the MER (maximum expected return) portfolio, which contains *Nuclear* exclusively. If it wished to minimize risk, opting for the MV (minimum variance) portfolio, then a mix of 56 percent *Nuclear* and 44 percent *Coal* would be optimal. Therefore, the degree of risk aversion characterizing the United States

Figure 5.2: Efficient Electricity Portfolios for the United States (2003, SURE-based)



	ER	Risk		Coal	Nuclear	Gas	Oil	Wind
AP2003	5.00	3.10		56	21	18	3	2
MER	5.76	1.80			100			
MV	5.50	1.46		44	56			

	ER	Risk		Coal	Nuclear	Gas	Oil	Wind
AP2003	5.00	3.10		56	21	18	3	2
MER_C	5.20	2.32		60	25	10		5
MV_C	5.07	2.03		60	25	1	9	5

clearly matters. However, risk aversion has its price because opting for MV rather than MER would entail a productivity increase of 5.50 rather than 5.76 percent p.a. Still, the MV portfolio with its annual volatility of 1.46 percent beats the actual one whose productivity advance is 5.00 percent only, associated with an annual volatility of 3.10 percent. Yet a share of *Nuclear* amounting to 100 rather than 21 percent in the MER portfolio (56 rather than 21 percent in the MV portfolio, respectively) must be deemed unrealistic for the United States of 2003. Therefore, Figure 5.2B shows an efficient frontier that takes into account that the current portfolio could be adjusted at considerable cost only. Since adjustment costs are unknown, upper limits for $Coal \leq 60\%$, $Nuclear \leq 25\%$, $Oil \leq 10\%$, and $Wind \leq 5\%$ are imposed on the individual shares for simplicity to reflect technical feasibility.

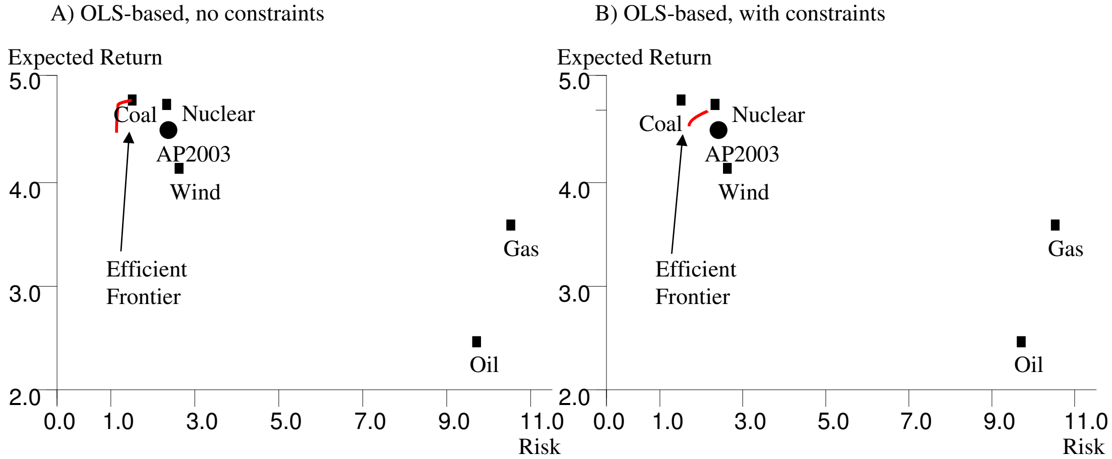
In the MER_C (with "C" for constrained) portfolio, the generation mix now contains 60 percent *Coal*, 25 percent *Nuclear*, 10 percent *Gas*, and 5 percent *Wind*, indicating that constraints are binding. Compared to the actual portfolio, productivity development would still speed up (from 5.00 percent p.a. to 5.20 percent p.a.), while volatility would be reduced from 3.10 to 2.32 percent p.a. In the MV_C alternative, the highest share is again allocated to *Coal* (60 percent, binding[10], up from 56 percent in the actual portfolio), followed by *Nuclear* (25 percent, binding, up from 21 percent), *Oil* (9 percent, up from 3 percent), and *Wind* (5 percent, again binding, up from 2 percent). The only technology to lose market share is *Gas* (a mere 1 percent, down from 18 percent). The rate of productivity increase would still attain 5.07 percent p.a. rather than 5.00 as in the actual portfolio, while risk declines to 2.03 from 3.10. One explanation of why *Gas* is almost phased out is its weak diversification effect, the correlation coefficient of its SURE residuals with *Coal* attaining 0.7605, the maximum value of panel A of Table 5.4. In the whole, current U.S. power generation is inefficient. It could be made more efficient by substituting *Gas* by *Coal*, *Nuclear*, *Oil* (not in the MER_C portfolio), and *Wind*.

If correlated shocks affecting generation costs would not have been taken into account (as in past studies), the results would have been very different, quite possibly misguiding the choice of an optimal technology mix. Figure 5.3 displays the OLS-based efficient frontiers and allocations. Without constraints (Figure 5.3A), the MER portfolio would have contained 100 percent *Coal*[11] (rather than 100 percent *Nuclear* as in the SURE-based case, see Figure 5.2A). The MV alternative, on the other hand, would have called for a portfolio with 63 percent *Coal*, 27 percent *Nuclear*, and 10 percent *Wind*, quite different from the SURE-based solution that excludes *Wind* while allocating 56 percent (rather than 27 percent) to *Nuclear*. Moreover, the United States would have little incentive to adjust its technology mix because OLS-based expected returns are only slightly higher and

10 Using portfolio theory for three U.S. generating technologies, Berger et al. (2003) also concluded that *Coal* dominates the MV portfolio with a share of 77 percent.

11 Berger et al. (2003), who do not control for correlation between unobserved shocks, also arrive at 100 percent *Coal*.

Figure 5.3: Efficient Electricity Portfolios for the United States (2003, OLS-based)



	ER	Risk	Coal	Nuclear	Gas	Oil	Wind
AP2003	4.50	2.40	56	21	18	3	2
MER	4.80	1.50	100				
MV	4.72	1.35	63	27			10

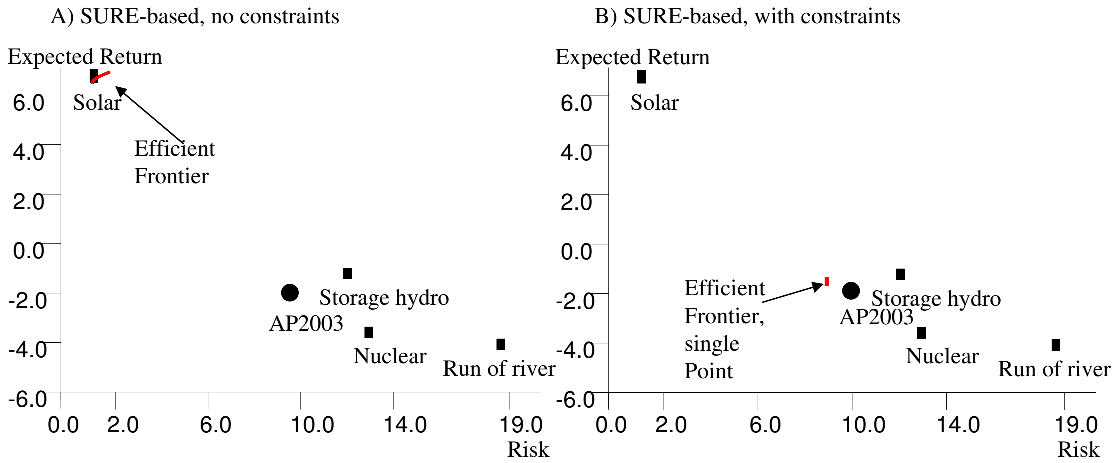
	ER	Risk	Coal	Nuclear	Gas	Oil	Wind
AP2003	5.00	3.10	56	21	18	3	2
MER_C	4.63	1.91	60	25	10		5
MV_C	4.55	1.70	60	25	2	8	5

volatilities slightly lower than estimated actual ones. With constraints imposed, however, OLS-based estimates would have resulted in efficient portfolios that practically coincide with the SURE-based ones (compare weights below Figures 5.2B and 5.3B). This was to be expected since most constraints are binding in both alternatives.

5.4.4 Efficient Provision of Swiss Electricity

Figure 5.4 displays the efficient electricity portfolios (as of 2003) for Switzerland. In Figure 5.4A (no constraints imposed), it is *Solar* rather than *Nuclear* (contrary to the United States) that dominates the MER portfolio with a 100 percent share. Opting for the MER portfolio, the country would achieve a productivity increase of 6.67 percent p.a. (rather than the 2 percent decrease with the actual portfolio), with volatility down from 10 to 1.05 percent p.a.. The MV portfolio consists of 98 percent *Solar* and 2 percent *Nuclear*, expected return being 6.43 percent p.a. and risk, a mere 1 percent. Clearly, in both countries non-CO2 emitting technologies (*Nuclear* in the United States and *Solar* in Switzerland) play a dominant role in the unconstrained efficient portfolios. However, shares of *Solar* close to 100 percent must be deemed unrealistic for Switzerland. Therefore, *Storage hydro*, *Run of river*, and *Solar* are constrained to their actual shares in 2003 (32, 24 and 4 percent p.a., respectively), leaving only *Nuclear* unconstrained. This can be justified by noting that *Storage hydro* and *Run of river* are already being utilized to full capacity (Laufer et al., 2004), while a share of *Solar* electricity of 4 percent constitutes the limit of what could have been achieved. The corresponding efficient frontier is shown in Figure 5.4B.

Figure 5.4: Efficient Electricity Portfolios for Switzerland (2003, SURE-based)



	ER	Risk		Nuclear	Storage hydro	Run or river	Solar
AP2003	-2.00	10.00		40	32	24	4
MER	6.67	1.05					100
MV	6.43	1.00		2			98

	ER	Risk		Nuclear	Storage hydro	Run or river	Solar
AP2003	-2.00	10.00		40	32	24	4
MER_C	-1.82	8.89		64	32		4
MV_C	-1.82	8.89		64	32		4

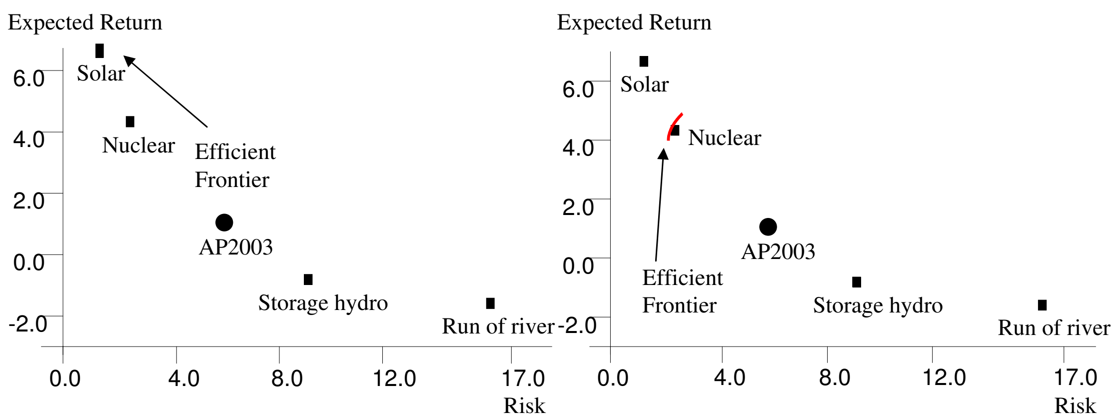
The MER_C portfolio calls for a complete substitution of *Run of river* (actual share 24 percent) by *Nuclear* (64 percent), *Storage hydro* (32 percent, binding), and *Solar* (4 percent, binding). This surprising result is due to the fact that *Run of river* is highly correlated with *Storage hydro*, indicating that it has no diversification potential (see the correlation coefficient of 0.5054 in Table 5.5A). At the same time, this technology has been subject to productivity decrease (see panel A of Table 5.3). In all, Figure 5.4B suggests that if "realistic" constraints are respected, Swiss power generation could be made more efficient by allowing the share of *Nuclear* to substantially increase and abandoning *Run of river*. Expected return would slightly increase, from -2.00 (actual) to -1.82 percent, regardless of choice between MER and MV portfolios, and volatility would drop from 10 (actual) to 8.89.

Results based on OLS-estimated efficient frontier are displayed in Figure 5.5. Acting on OLS-based estimates, Switzerland would have expected marked productivity increases rather than the decreases implied by SURE, at the same time severely underestimating volatility. Finally, the country would have wrongly slashed the share of *Storage hydro* from 32 percent to 0 percent (MER_C) or 8 percent (MV_C), respectively. Therefore, the choice of statistical specification may again well matter for decision-making.

Figure 5.5: Efficient Electricity Portfolios for Switzerland (2003, OLS-based)

A) OLS-based, no constraints

B) OLS-based, with constraints



	ER	Risk		Nuclear	Storage hydro	Run or river	Solar
AP2003	1.00	6.00		40	32	24	4
MER	6.67	1.00					100
MV	6.67	1.00					100

	ER	Risk		Nuclear	Storage hydro	Run or river	Solar
AP2003	1.00	6.00		40	32	24	4
MER_C	4.41	2.12		96			4
MV_C	3.98	1.95		88	8		4

5.4.5 United States and Switzerland Compared

This section is devoted to a comparison of results obtained for the two countries as of the year 2003, using SURE-based estimates. Starting with no constraints imposed (Figures 5.2A and 5.4A), the United States could have achieved an average productivity increase of 5.76 p.a. by adopting the MER portfolio, Switzerland even 6.67 percent p.a. However, both countries would have had to completely change the composition of their technology portfolio, to 100 percent *Nuclear* (United States) and 100 percent *Solar* (Switzerland), respectively. Turning to the MV alternative, the volatility reduction achieved amounts to 1.54 percentage points (3.10 - 1.46) for the United States, much less than for Switzerland with its 9.00 percentage points (10.00 - 1.00). The implications in terms of portfolio composition are quite different for the two countries as well. Whereas opting for the MV alternative calls for 56 percent (rather than 100 percent) *Nuclear* in the case of the United States, it would leave *Solar* at almost 100 percent in the case of Switzerland. Since shares close to 100 percent are far from reality in either country, constraints on admissible shares of technologies were imposed in Figures 5.2B and 5.4B. This causes the existing amount of diversification to diminish in both countries, with *Coal* (United States) and *Nuclear* (Switzerland) becoming the principal energy sources. However, only the Swiss expected rate of return drops (from a 6.43 percent productivity increase to a 1.82 percent decrease for the MV_C portfolio), associated with a marked surge in volatility.

On the whole, it appears that the U.S. electricity industry, while respecting feasibility constraints, would have gained by substituting *Gas* by *Coal*, *Nuclear*, and *Wind* technologies by 2003, regardless of the choice between the MER_C and the MV_ C portfolio. Swiss utilities would have stood to gain as well by adopting more *Nuclear* to the detriment of *Run of river*, an important source of primary energy until recently. Divergences of U.S. and Swiss actual choices and efficient choices likely arose in the past since generating technologies have been selected solely on an individual, case-per-case basis, failing to consider their contribution to overall portfolio performance. Both industries at present fall short of their respective efficiency frontiers. In the United States, the gap amounts to a foregone 0.07 to 0.20 percentage points productivity increase p.a. and 0.78 to 1.07 points reduction of volatility (see Figure 5.2B). In Switzerland, the estimates amount to a foregone 0.18 percentage points of productivity growth and 1.11 points reduction of risk (see Figure 5.4B). Therefore, there is some evidence suggesting that the more heavily regulated Swiss industry is characterized by a higher degree of inefficiency in the allocation of generating technologies than its largely deregulated U.S. counterpart.

5.5 Conclusion

The objective of this contribution is to determine the efficient provision of electricity in the United States (traditionally fossil-based) and Switzerland (traditionally hydro- and nuclear-based), applying portfolio theory. The observation period covers 1982 to 2003 (United States) and 1986 to 2003 (Switzerland), respectively. Because the error terms proved to be correlated across equations explaining technology-specific productivity changes, Seemingly Unrelated Regression Estimation (SURE) was adopted for estimating the covariance matrix used in determining the efficient frontier. Interestingly, the maximum expected return (MER) portfolios of both countries boil down to one non-CO2 energy source (*Nuclear* in the United States and *Solar* in Switzerland). When constraints limiting changes from the status quo are imposed to reflect the high cost associated with adjusting the technology mix, the MER_C portfolio for the United States contains 60 percent *Coal* (up from 56 percent) and for Switzerland, 64 percent *Nuclear* (up from 40 percent).

However, one could argue that for populations as risk-averse as the American and the Swiss (Szpiro, 1986), the minimum variance portfolio (MV) is appropriate. Adopting the MV criterion and imposing the same constraints, U.S. utilities would still want to assign 60 percent of their portfolio to *Coal*, almost entirely replacing *Gas*. The productivity changes and hence returns of *Gas* are not only highly volatile but also strongly correlated with those of other technologies, depriving it of a possible diversification effect. At the same time, *Coal*-generated electricity became cleaner, causing (initially high) external costs to fall and making *Coal* very attractive from an investor point of view. In the Swiss MV_C portfolio (subject to constraints), *Nuclear* accounts for even 64 percent while *Run of river* drops out (down from 24 percent). One is therefore led to conclude that as of 2003, both the U.S. and Swiss technology mix are inefficient even if "realistic" constraints are respected. While U.S. utilities are currently closer to their efficiency frontier than their more heavily regulated Swiss counterparts, they still may reap efficiency gains by investing more in *Coal* and moving away from *Gas*.

The choice of econometric methodology proves important for decision-making. Efficiency frontiers estimated by OLS would tend to underestimate both expected returns and risk reduction potential in the case of the United States but overestimate achievable expected returns and underestimating risk reduction in the case of Switzerland. These discrepancies largely vanish, however, when feasibility constraints are imposed. Still, failure to account for correlation between unobserved shocks impinging on the different generation technologies using SURE does create the risk of opting for an inefficient solution. This finding contrasts with Berger et al. (2003), who concluded that the outcome of portfolio analysis is insensitive to econometric estimation techniques. However, the present study agrees with earlier ones in suggesting that utilities and policy makers, by adopting a single-technology approach typical of conventional least-cost planning, fail to take account of correlations between risky generating technologies. The consequence is a portfolio of generating technologies that is inefficient, achieving a too low expected rate of return and/or suffering from excessive volatility.

These statements are based on an investor view. One could also adopt a current user view, which emphasizes productivity levels rather than productivity changes over time. Future contributions therefore may compare the two views. They could also focus on prediction rather than postdiction, examining whether emergent new technologies are part of future efficient frontiers. Finally, the strong assumption of a once-and-for-all decision regarding the choice of technology needs to be relaxed. A real options approach could be used to account for the irreversibility often inherent in the decision to adopt a technology. Deferring adoption may become the preferred choice in the face of stochastic productivity changes caused e.g. by a liberalization of electricity markets – or its failure to materialize as expected. Still, the present study provides first indications of where to go in the future in an attempt to reach the efficient mix of power-generating technologies in countries that are as diverse as e.g. the United States and Switzerland.

CHAPTER VI

Conclusion

6

Conclusion

This chapter summarizes the key findings and policy recommendations of each essay, followed by possible suggestions for future research.

The objective in Chapter 2 was to determine the efficiency of all 26 member states (cantons) in Switzerland and examine whether fiscal equalization influences their efficiency. Aggregate output performance indicators, including six major public services, were constructed to calculate cantonal efficiency scores based on robust data envelopment analysis. The results suggest two main policy implications. First, a comparison of the six service categories revealed that cantons with high overall performance do not automatically outperform in all categories, preventing any one from becoming dominant in Tiebout competition. Second, the equity-efficiency trade-off noted by Stiglitz (1988) seems to exist in Switzerland. Both cantons that are payers and cantons that are receivers seek to keep their 'financial potential' low by producing public services at higher than minimum cost, the former to ease their burden and the latter to receive more transfer payments and subsidies. However, substituting earmarked federal subsidies, the main component of transfer payments prior to 2008, for freely disposable lump-sum subsidies as part of the new equalization program implemented in 2008 is likely to enhance cantonal performance.

The results presented in Chapter 3 support the relevance of accounting for heterogeneity in the measurement of hospital performance. Standard frontier models, which assume homogeneous technology for all hospitals, fail to detect all the cost variability among hospitals that is caused by heterogeneity due to exogenous influences. Efficiency estimates of approximately 100 Swiss hospitals for the years 2004 to 2007 reveal that those rated 85 percent efficient or less (using the standard frontier model) would gain up to 12 percentage points when unobserved heterogeneity is taken into account using a random parameter frontier model. Therefore, quite a few highly efficient hospitals could end up in financial distress if regulators cut reimbursement rates in an attempt to enforce the cost reductions indicated by the standard frontier model.

Chapter 4 provides a confirmation of the theoretical expectation that a prospective payment system increases hospital cost efficiency. Payment systems designed to put hospitals at operating risk seem to be more effective at reducing hospital costs than retrospective payment systems. However, with respect to the Swiss hospital payment reform effective in 2012, this requires that the implementation be fully prospective and must preclude any bailouts. Results relating to the global budget system reveal that if hospitals can obtain higher budgets to cover past errors, then the incentive for cost minimization disappears. In addition, the settings for the remuneration per admission are also important. Whereas general remuneration settings with a per diem element can be used to unnecessarily keep a patient in the hospital, a DRG system strengthens incentives for cost minimization. However, DRG is not fully effective after initiation and additional efficiency gains occur later on, but these are smaller in the third year than in the second year.

Finally, in Chapter 5 a portfolio theory for a risk adjusted efficiency measurement was applied to determine the efficient provision of electricity in the United States and Switzerland. The seemingly unrelated regression estimation (SURE) method was adopted for estimating the covariance matrix used in determining the efficient frontier. The results support the notion that adopting a single-technology approach, which is typical of conventional least-cost planning, fails to take into account correlations between risky generating technologies. In particular, populations as risk-averse as Americans (Szpiro, 1986) are strongly advised to adopt a feasible minimum variance portfolio comprised of 60 percent coal, 25 percent nuclear, 9 percent oil, 5 percent wind, and 1 percent gas. In the Swiss case, a feasible portfolio for a risk-averse population would include 64 percent nuclear, 32 percent storage hydro, and 4 percent solar. Run-of-the-river (ROR) hydroelectricity drops from 24 percent in the status quo to zero in the minimum variance portfolio, mostly due to diversification effects. In regards to their actual 2003 portfolios, both the U.S. and Swiss technology mixes were inefficient even if “realistic” constraints are applied. While U.S. utilities are currently closer to their efficiency frontier than their more strictly regulated Swiss counterparts, they may still reap efficiency gains by investing more in coal and moving away from gas.

This thesis has spotlighted four areas where efficiency measurement techniques are useful in uncovering market inefficiency, providing potential targets for cost reduction. Switzerland could increase its competitiveness with policy reforms that increase the incentive for efficient transformation of production inputs into outputs. This is not necessarily limited to the provision of public goods or to the hospital and electricity markets. It could be an indication for all markets where conditions for Pareto efficiency are restricted. Therefore, although all the results presented in Chapters 2 through 5 appear to be very specific, they address an issue of general importance and provide starting points for further research. Here, four possible extensions are summarized.

The first extension is methodological and aims to incorporate the production uncertainty in Chapter 5 into the single-technology framework of Chapters 2 through 4. Although many articles have already examined the development of production technologies over time, no study has incorporated technology uncertainty in the estimates. Technology change is always measured deterministically, meaning that the entire distance between the technology frontiers is measured as technology improvement. A frequently applied measurement instrument is the Malmquist index introduced by Caves et al. (1982). However, as shown in Chapter 5, exogenous shocks can influence productivity as well, making productivity changes over time rather stochastic. Thus, the distance between the technology frontiers should be split into a productivity and a random noise term, congruent to the stochastic frontier analysis for individual observations (see Chapters 3 and 4). Therefore, the model of Chapter 5 could be extended with a composite error term that incorporates a positively distributed productivity term and a normally distributed random noise term for technology risk. The application of an extended model would not only enable unbiased estimates of technology improvement, it would also reveal information about the importance of market uncertainty.

The second extension is related to the recent merging of local communities in several member states (e.g., Luzern, Glarus) with the goal of achieving cost efficiency gains. Politicians expect the potential for economies of scale in the provision of public goods. Although Chapter 2 reveals some evidence in favor of decentralization, no detailed results are shown regarding economies of scale or the optimal size of local communities. This is an issue for further research. However, the optimal size of a local community could differ among public services. Thus, aggregation on a local community level may be not efficient (Frey, 2005).

Chapter 3 and 4 are also valuable starting points for future research. Although cost savings have been found for hospitals with a prospective payment system, the results are not yet sufficient to make any conclusion about its influence on health care expenditure. In particular, health care expenditure may increase if cost savings are only due to a cost shift from inpatient to outpatient or rehabilitation care or if the number of admissions increases. Further research could address these situations.

Finally, the case of Switzerland makes it clear that the assumption of an autarky market in Chapter 5 is not realistic. In 2009, Switzerland exported 54.2 billion kWh of the 66.5 billion kWh generated domestically and simultaneously imported 52 billion kWh (see BFE, 2009). Thus, it would be more realistic to use an efficiency measure for an open market, where imports and exports are used as an alternative to domestic production. This is of interest for further study as well.

Bibliography

Afonso, A. and S. Fernandes (2006). Measuring local government spending efficiency: Evidence for the Lisbon region. *Regional Studies 40*(1), 39–53.

Afonso, A. and C. Scaglioni (2005). Public services efficiency provision in Italian regions: a non-parametric analysis. Working Paper 02, Department of Economics at the School of Economics and Management (ISEG), Technical University of Lisbon.

Afonso, A., L. Schuknecht, and V. Tanzi (2006). Public sector efficiency: Evidence for new EU member states and emerging markets. Working Paper 01, Department of Economics at the School of Economics and Management (ISEG), Technical University of Lisbon.

Aigner, D., C. Lovell, and P. Schmidt (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics 6*(1), 21–37.

Al-Subaihi, A. A. (2002). Variable selection in multivariable regression using sas/iml. *Journal of Statistical Software 07*(i12).

Andersen, P. and N. C. Petersen (1993). A procedure for ranking efficient units in data envelopment analysis. *Management Science 39*(10), 1261–1264.

Awerbuch, S. (2005). The cost of geothermal energy in the western US regions: A portfolio based approach. Sandia report, Sandia National Laboratories, TN, USA.

Awerbuch, S. (2006). Portfolio-based electricity generation planning: Policy implications for renewables and energy security. *Mitigation and Adaptation Strategies for Global Change 11*, 693–710.

Awerbuch, S. and M. Berger (2003). Energy security and diversity in the EU: A mean-variance portfolio approach. IEA Report EET/2003/03.

Bar-Lev, D. and S. Katz (1976). A portfolio approach to fossil fuel procurement in the electric utility industry. *Journal of Finance 31*(3), 933–947.

Barankay, I. and B. Lockwood (2007). Decentralization and the productive efficiency of government: Evidence from swiss cantons. *Journal of Public Economics 91*(5-6), 1197–1218.

Battese, G. E. and T. J. Coelli (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics 20*(2), 325–332.

Berger, M., S. Awerbuch, and R. Haas (2003). Versorgungssicherheit und Diversifizierung der Energieversorgung in der EU (Security of supply and diversification of energy supply in the EU). Report, Bundesamt für Verkehr, Innovation und Technologie, Wien (Federal Office for Transportation, Innovation and Technology, Vienna).

BFE (2009). Schweizerische Elektrizitätsstatistik 2009 (Electricity statistics of Switzerland 2009). Report, Bundesamt für Energie, (Federal Office for Energy).

Biorn, E., T. Hagen, T. Iversen, and J. Magnussen (2003). The effect of activity-based financing on hospital efficiency: A panel data analysis of DEA efficiency scores 1992-2000. *Health Care Management Science 6*, 271–283.

Biorn, E., T. P. Hagen, T. Iversen, and J. Magnussen (2006). Heterogeneity in hospitals' responses to a financial reform: A random coefficient analysis of the impact of activity-based financing on efficiency. Working Paper 8100, University Library of Munich, Germany.

Bodie, Z., A. Kane, and A. J. Marcus (2005). *Investments* (6 ed.). New York: McGraw-Hill.

Borden, J. P. (1988). An assessment of the impact of diagnosis-related group (DRG)-based reimbursement on the technical efficiency of new jersey hospitals using data envelopment analysis. *Journal of Accounting and Public Policy 7*(2), 77–96.

Caves, D. W., L. R. Christensen, and W. E. Diewert (1982). Multilateral comparisons of output, input, and productivity using superlative index numbers. *Economic Journal 92*(365), 73–86.

Chalkley, M. and J. M. Malcomson (1998). Contracting for health services when patient demand does not reflect quality. *Journal of Health Economics 17*(1), 1–19.

Chalkley, M. and J. M. Malcomson (2000). Government purchasing for health services. In A. J. Culyer and J. P. Newhouse (Eds.), *Handbook of Health Economics*, Volume 1A, pp. 847–889. Amsterdam: North Holand: Elsevier Science Ltd.

Chern, J.-Y. and T. T. H. Wan (2000). The impact of the prospective payment system on the technical efficiency of hospitals. *Journal of Medical Systems 24*, 159–172.

Coelli, T., D. S. P. Rao, C. J. O'Donnell, and G. E. Battese (2005). *An Introduction to Efficiency and Productivity Analysis*. USA: Springer.

De Borger, B. and K. Kerstens (1996). Cost efficiency of Belgian local governments: A comparative analysis of FDH, DEA, and econometric approaches. *Regional Science and Urban Economics 26*(2), 145–170.

Debreu, G. (1951). The coefficient of resource utilization. *Econometrica 19*(3), 273–292.

Diewart, W. E. and A. O. Nakamura (1999). Benchmarking and the measurement of best practice efficiency: An electricity generation application. *The Canadian Journal of Economics / Revue canadienne d'Economique 32*(2), 570–588.

Dismuke, C. and V. Sena (1999). Has the DRG system influenced the efficiency of diagnostic technology in Portugal? an empirical analysis using parametric and non-parametric methods. *Health care management science 2*, 107–116.

Drake, L. and R. Simper (2003). The measurement of English and Welsh police force efficiency: A comparison of distance function models. *European Journal of Operational Research 147*(1), 165–186.

EMG (2000). Elektrizitätsmarktgesetz (Law on electricity market). http://www.admin.ch.

EMV (2002). Elektrizitätsmarktverordnung (Ornance on electricity markets). http://www.admin.ch.

European Commission (2003). External costs. Report, European Commission (EC).

Farrell, M. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Scociety 120*(3), 253–281.

Farsi, M., A. Fetz, and M. Filippini (2008). Economies of scale and scope in multi-utilities. *The Energy Journal 29*(4), 123–144.

Farsi, M. and M. Filippini (2006). An analysis of efficiency and productivity in Swiss hospitals. *Swiss Journal of Economics and Statistics (SJES) 142*(I), 1–37.

Farsi, M., M. Filippini, and W. Greene (2006). Application of panel data models in benchmarking analysis of the electricity distribution sector. *Annals of Public and Cooperative Economics 77*(3), 271–290.

Feld, L. and J. Matsusaka (2003). Budget referendums and government spending: Evidence from Swiss cantons. *Journal of Public Economics 87*(12), 2703–2724.

Filippini, M., M. Farsi, L. Crivelli, and M. Zola (2004). An analysis of efficiency and productivity in Swiss hospitals. Report, Swiss Federal Statistical Office.

Fischer, R. (2004). Die Unterschiede in der Steuerbelastung der Kantone: Eine Analyse auf der Basis eines mikroökonomischen Haushaltsmodells (Differences in cantonal tax burdens: alysis based on a microeconomic houshold model). Working Paper 06, Federal Administration of Switzerland.

Fischer, R., T. Beljean, and J. Fivaz (2003). Mehr Chancengleichheit und Eigenverantwortung: der neue Schweizer Finanzausgleich zwischen Bund und Kantonen (More equality in opportunities and responsibility: the new fiscal equalization scheme between the confederation and the cantons). *Vierteljahreshefte zur Wirtschaftsforschung 72*(3), 407–422.

Frey, B. S. (2005). Functional, overlapping, competing jurisdictions: Redrawing the geographic borders of administration. *European Journal of Law Reform 3*(4), 543–555.

Frey, R., B. Dafflon, C. Jeanrenaud, and A. Meier (1994). Der Finanzausgleich zwischen Bund und Kantonen (Fiscal equalization between the confederation and the cantons). Technical report, Federal Administration of Switzerland.

Gantner, U., M. Jakob, and S. Hirschberg (2000). Perspektiven der zukünftigen Energieversorgung in der Schweiz unter Berücksichtigung von nachfrageorientierten Massnahmen (Perspectives on the future provision of energy in Switzerland, with special emphasis on demand-side-management). Report, Switzerland: Paul Scherrer Institute.

Greene, W. (2004a). Distinguishing between heterogeneity and inefficiency: stochastic frontier analysis of the world health organization's panel data on national health care systems. *Health Economics 13*(10), 959–980.

Greene, W. (2004b). Interpreting estimated parameters and measuring individual heterogeneity in random coefficient models. Working Paper 04-08, New York University, Leonard N. Stern School of Business, Department of Economics.

Greene, W. (2005a). Fixed and random effects in stochastic frontier models. *Journal of Productivity Analysis 23*, 7–32.

Greene, W. (2005b). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics 126*(2), 269–303.

Griffin, J. and M. Steel (2007). Bayesian stochastic frontier analysis using WinBUGS. *Journal of Productivity Analysis 27*, 163–176.

Grossman, P. J., P. Mavros, and R. W. Wassmer (1999). Public sector technical inefficiency in large U.S. cities. *Journal of Urban Economics 46*(2), 278–299.

Hirschberg, S. and M. Jakob (1999). Cost structure of the Swiss electricity generation under consideration of external costs. Tagungsband (conference volume), SAEE Seminar, Bern.

Hollingsworth, B. (2008). The measurement of efficiency and productivity of health care delivery. *Health Economics 17*(10), 1107–1128.

Huang, C. and J. Liu (1994). Estimation of a non-neutral stochastic frontier production function. *Journal of Productivity Analysis 5*, 68–80.

Huang, H.-C. (2004). Estimation of technical inefficiencies with heterogeneous technologies. *Journal of Productivity Analysis 21*(3), 277–296.

Humphreys, H. B. and K. T. McClain (1998). Reducing the impacts of energy price volatility through dynamic portfolio selection. *The Energy Journal 19*(3), 107–132.

IAE (2005). Key World Energy Statistics 2004. Report, Paris: International Energy Agency.

Inman, R. P. (2008). The flypaper effect. Working Paper 14579, National Bureau of Economic Research.

Jacobs, R., P. C. Smith, and A. Street (2006). *Measuring Efficiency in Health Care.* Cambridge: Cambridge University Press.

Keeler, E. B. (1990). What proportion of hospital cost differences is justifiable? *Journal of Health Economics 9*(3), 359–365.

Kienzle, F. and G. Andersson (2008). Efficient multi-energy generation portfolios for the future. Working paper, ETH Zurich, Power Systems Laboratory.

KKG (2005). Annual report. www.kkg.ch.

KKl (2005). Medienkonferenz 20 Jahre KKL, 10. Januar 2005 Portrait - Fakten - Zahlen zu 20 Jahre Kernkraftwerk Leibstadt (Portrait, facts and figures concerning 20 years of the nuclear plant at Leibstadt). www.kkl.ch.

Knott, G. (2000). *Interpolating Cubic Splines.* Boston: Birkhäuser.

Koop, G., J. Osiewalski, and M. F. J. Steel (1997). Bayesian efficiency analysis through individual effects: Hospital cost frontiers. *Journal of Econometrics 76*(1-2), 77–105.

Koopmans, T. (1951). An analysis of production as an efficient combination of activities. In T. Koopmans (Ed.), *Activity Analysis of Production and Allocation*, pp. 33–97. New York: John-Wiley and Sons, Inc.

Krey, B. (2008). Scope of electricity efficiency improvement in Switzerland until 2035. Working Paper 0813, University of Zurich, Socioeconomic Institute.

Krey, B. and P. Zweifel (2009). Efficient and secure power for the United States and Switzerland. In M. Bazilian and F. Roques (Eds.), *Analytical Methods for Energy Diversity - Mean-Variance Optimization for Electric Utilities*, Energy Policy and Economics Series, pp. 193–218. Elsevier.

Kumar, S. and S. Gupta (2004). Resource use efficiency of US electricity generating plants during the SO2 trading regime: A distance function approach. Working Paper 04/17, National Institute of Public Finance and Policy.

Kumbhakar, S. and C. A. K. Lovell (2000). *Stochastic frontier analysis: an econometric approach.* Cambridge: Cambridge University Press.

Kumbhakar, S. C., S. Ghosh, and J. T. McGuckin (1991). A generalized production frontier approach for estimating determinants of inefficiency in U.S. dairy farms. *Journal of Business & Economic Statistics 9*(3), 279–286.

Laufer, F., S. Grötzinger, M. Peter, and A. Schmutz (2004). Ausbaupotentiale der Wasserkraft (Potential for expansion of hydro power). Report, Bundesamt für Energie (Federal Office of Energy).

Liew, V. (2004). *On Autoregressive Order Selection Criteria.* Putra University Malysia.

Linna, M. (1999). The impact of health care financing reform on the productivity change in finnish hospitals. *System Analysis Laboratory Research Reports 79*(A).

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance 7*(1), 77–91.

Meeusen, W. and J. van den Broeck (1977). Efficiency estimation from cobb-douglas production functions with composed error. *International Economic Review 18*(2), 435–444.

Meister, U. (2008). Spitäler zwischen Politik und Wettbewerb: Betriebliche Autonomie im Kantonsvergleich(hospitals between regulation and competition). Report, Avenir Suisse.

Morey, R. and D. Dittman (1996). Cost pass-through reimbursement to hospitals and their impacts on operating efficiencies. *Annals of Operations Research 67*, 117–139.

NEPG (2004). *U.S. National Energy Policy Report. Report of the National Energy Policy Development Group.* U.S. Government Printing Office.

Newhouse, J. P. (1996). Reimbursing health plans and health providers: Efficiency in production versus selection. *Journal of Economic Literature 34*(3), 1236–1263.

Ntzoufras, I. (2009). *Bayesian modeling using WinBUGS.* Wiley Series in Computational Statistics. Hoboken, NJ: Wiley.

Olson, M. J. (1969). The principle of fiscal equivalence: The division of responsibilities among different levels of government. *The American Economic Review 59*(2), 479–487.

Orea, L. and S. C. Kumbhakar (2004). Efficiency measurement using a latent class stochastic frontier model. *Empirical Economics 29*(1), 169–183.

Pareto, V. (1897). *Cours d'économie politique.* Band 2. Lausanne: F. Rouge.

Pitt, M. M. and L.-F. Lee (1981). The measurement and sources of technical inefficiency in the indonesian weaving industry. *Journal of Development Economics 9*(1), 43–64.

Pommerehne, W. and P. Zweifel (1991). Success of a tax amnesty: At the polls, for the fisc? *Public Choice 72*(2-3), 131–165.

Resende, M. (2002). Relative efficiency measurement and prospects for yardstick competition in Brazilian electricity distribution. *Energy Policy 30*(8), 637–647.

RWE (2005). Data on solar generated electricity. www.rewschottscolar.com.

Schmidt, P. and R. C. Sickles (1984). Production frontiers and panel data. *Journal of Business & Economic Statistics 2*(4), 367–374.

Schneider, P. (2007). Provider payment reforms: lessons from Europe and America for South Eastern Europe. Discussion paper, World Bank, Washington.

Seiford, L. and J. Zhu (2002). Modeling undesirable factors in efficiency evaluation. *European Journal of Operational Research 142*, 16–20.

Shephard, R. (1970). *Theory of Cost and Production Functions*. Princeton: Princeton University Press.

Simar, L. and P. W. Wilson (2000). Statistical inference in nonparametric frontier models: The state of the art. *Journal of Productivity Analysis 13*, 49–78.

Smith, B. (1985). *Decentralization: the Territorial Dimension of State*. London: George Allen & Unwin.

Smith, P. C. (2004). Health care reforms in Europe and their implications for Japan. *The Japanese Journal of Social Security Policy 3*(2), 80–95.

Sommersguter-Reichmann, M. (2000). The impact of the Austrian hospital financing reform on hospital productivity: empirical evidence on efficiency and technology changes using a non-parametric input-based malmquist approach. *Health Care Management Science 3*(4), 309–321.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B 64*, 583–640.

Steinmann, L., G. Dittrich, A. Karmann, and P. Zweifel (2004). Measuring and comparing the (in)efficiency of German and Swiss hospitals. *European Journal of Health Economics 5*(3), 216–226.

Steinmann, L. and P. Zweifel (2003). On the (in)efficiency of Swiss hospitals. *Applied Economics 35*(3), 361–370.

Stiglitz, J. (1988). *Economics of the Public Sector*. USA: Norton & Company.

Stiglitz, J. E. (1991). The invisible hand and modern welfare economics. NBER Working Paper 3641, National Bureau of Economic Research, Inc.

Stutzer, A. (1999). A democracy index for the Swiss cantons. Working Paper 23, Institute for Empirical Research in Economics, University of Zurich.

Swiss Federal Office of Energy (SFOE) (2003). Schweizerische Elektrizitätsstatistik 2003 (Swiss electricity statistics 2003). Report, Bundesamt für Energie (Federal Office of Energy).

Szpiro, G. G. (1986). Über das Risikoverhalten in der Schweiz (On risk behavior in Switzerland). *Swiss Journal of Economics and Statistics (SJES) 122*(III), 463–470.

Thanssoulis, E. (1999). Setting achievement targets for school children. *Education Economics 7*(2), 101–119.

Thöny, B. (2005). *Der Finanzausgleich im Wettbewerbsföderalismus - Eine Theorie des Finanzausgleichs und eine Analyse der innerkantonalen Finanzausgleichssysteme in der Schweiz (Fiscal equalization in competitive federalism – Theory of fiscal equalization and analysis of intercantonal systems of fiscal equalization in Switzerland).* Bamberg: Difo Druck GmbH.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy 64*, 416–424.

Tsionas, E. G. (2002). Stochastic frontier models with random coefficients. *Journal of Applied Econometrics 17*(2), 127–147.

UIC (2005). The economics of nuclear power. http://www.uic.com.au /nip08.htm.

Vanden Eeckaut, P., H. Tulkens, and M. Jamar (1993). Cost efficiency in Belgian municipalities. In H. Fried, C. Lovell, and S. Schmidt (Eds.), *The Measurement of Productive Efficiency: Techniques and Applications*, pp. 300–334. New York: Oxford University Press.

Widmer, P. K. (2011). Does prospective payment increase hospital (in)efficiency? evidence from the Swiss hospital sector. Working Paper 53, Department of Economics, University of Zürich.

Widmer, P. K., P. Zweifel, and M. Farsi (2011). Accounting for heterogeneity in the measurement of hospital performance. Working Paper 52, University of Zurich, Department of Economics.

Wooldridge, J. (2003). *Introductory Econometrics: A Modern Approach* (2 ed.). Thomson South-Western.

Worthington, A. (2004). Frontier efficiency measurement in healthcare: A review of empirical techniques and selected applications. *Medical Care Research and Review 61*(2), 1–36.

Worthington, A. C. (2001). An empirical survey of frontier efficiency measurement techniques in education. *Education Economics 9*(3), 245–268.

Worthington, A. C. and B. E. Dollery (2001). Measuring efficiency in local government: An analysis of new south wales municipalities' domestic waste management function. *Policy Studies Journal 29*(2), 232–250.

Yu, Z. (2003). A spatial mean-variance mip model for energy market risk analysis. *Energy Economics 25*(3), 255–268.

Zweifel, P. and R. Eisen (2011). *Insurance Economics*. Springer (forthcoming).

Curriculum Vitae

Philippe Widmer was born on May 26, 1982 in Cham near Zug (Switzerland). After completing his high school in Zug he studied Economics and Finance at the University of Zurich (Switzerland) and the University of Tulsa (USA). He graduated in 2007 at the University of Zurich with a diploma thesis in econometrics. From 2007 to 2011 he has been a research assistant at the chair of Prof. Dr. Peter Zweifel (Department of Economics, University of Zurich) and wrote his doctoral thesis in economics. During his time at the chair of Prof. Dr. Peter Zweifel he supervised various university courses, was bachelor and master student advisor, and worked as a consultant for various third-party funded projects. Since 2008, he is working in part-time employment at Polynomics AG, an economic consultancy firm.